NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

AN ANALYSIS OF “SHOOT-AND-SCOOT” TACTICS

by

Younglak Shim

March 2017

Thesis Advisor: Michael Atkinson
Second Reader: Dashi Singham

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** March 2017	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE AN ANALYSIS OF "SHOOT-AND-SCOOT" TACTICS	**5. FUNDING NUMBERS**
6. AUTHOR(S) Younglak Shim	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB number ____N/A____.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

Firing multiple rounds of artillery from the same location has several benefits: a high rate of fire at the enemy and potentially improved accuracy as the shooter's aim adjusts to previous rounds. Firing many rounds from the same location, however, carries significant risk that the enemy will detect the location of the artillery. Therefore, the shooter may want to periodically change location to avoid counter-battery fire. This maneuver is known as the shoot-and-scoot tactic. The importance of the shoot-and-scoot tactic has increased in recent years with the prevalence of self-propelled artillery and significant improvements in counter-detection technology such as radar. This thesis analyzes the shoot-and-scoot tactic using stochastic models, such as continuous-time Markov chains. We explore various examples and conclude that spending a reasonable amount of time firing multiple shots in the same location is preferable to moving immediately after firing one shot. Moving frequently reduces risk to artillery, but limits the artillery's ability to inflict damage on the enemy. These results should provide commanders with insight about how frequently they should change positions based on the risk level and their capabilities.

14. SUBJECT TERMS shoot-and-scoot, Markov chain, stochastic process, artillery, counter-battery radar, counter-fire	**15. NUMBER OF PAGES** 79
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

AN ANALYSIS OF "SHOOT-AND-SCOOT" TACTICS

Younglak Shim
Captain, Republic of Korea Army
B.S., Korea Military Academy, 2008

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 2017

Approved by: Michael Atkinson
Thesis Advisor

Dashi Singham
Second Reader

Patricia Jacobs
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Firing multiple rounds of artillery from the same location has several benefits: a high rate of fire at the enemy and potentially improved accuracy as the shooter's aim adjusts to previous rounds. Firing many rounds from the same location, however, carries significant risk that the enemy will detect the location of the artillery. Therefore, the shooter may want to periodically change location to avoid counter-battery fire. This maneuver is known as the shoot-and-scoot tactic. The importance of the shoot-and-scoot tactic has increased in recent years with the prevalence of self-propelled artillery and significant improvements in counter-detection technology such as radar. This thesis analyzes the shoot-and-scoot tactic using stochastic models, such as continuous-time Markov chains. We explore various examples and conclude that spending a reasonable amount of time firing multiple shots in the same location is preferable to moving immediately after firing one shot. Moving frequently reduces risk to artillery, but limits the artillery's ability to inflict damage on the enemy. These results should provide commanders with insight about how frequently they should change positions based on the risk level and their capabilities.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I. INTRODUCTION 1
   A. BACKGROUND 1
   B. MOTIVATION 1
   C. SCOPE 2
   D. LITERATURE REVIEW 3
   E. THESIS OUTLINE 5

II. LONG-RUN RISK MODEL 7
   A. CTMC REVIEW 7
   B. MODEL DESCRIPTION 9
   C. THE LONG-RUN BEHAVIOR 12
   D. GENERAL NUMBER OF RISK LEVELS 13
   E. OBJECTIVE FUNCTIONS FOR OPTIMIZATION 15
   F. NUMERICAL DEMONSTRATION 16
   G. REWARD RENEWAL PROCESS 26
      1. Deterministic T 26
      2. Deterministic M 29
   H. SUMMARY 31

III. WIN PROBABILITY MODEL WITH TIME LIMIT 33
   A. THE STATES 33
   B. MODEL DESCRIPTION 35
   C. WINNING PROBABILITY 39
   D. OPTIMIZATION 41
   E. NUMERICAL DEMONSTRATION 44
      1. Parameters 44
      2. Algorithms 45

IV. CONCLUSION 51
   A. SUMMARY 51
   B. FUTURE WORK 52

LIST OF REFERENCES 55

INITIAL DISTRIBUTION LIST 57

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

Figure 1. Transition diagram and its infinitesimal generator matrix Q 11
Figure 2. Advanced CTMC model 14
Figure 3. Plots using different n and change patterns of μ_i 19
Figure 4. Plots using the objective function Z_2 20
Figure 5. Examples of distribution of π_i , n = 10 21
Figure 6. Plots using the objective function Z_3 22
Figure 7. Plots when Red has a higher effective firing rate 24
Figure 8. Plots when Blue has a higher effective firing rate 25
Figure 9. Relative effective firing rate ($f_{B,Ri} - f_{R,Ri}$) in the cases when Blue has a higher firing rate 26
Figure 10. Long-run average reward rate by λ 28
Figure 11. Comparison with the CTMC model 29
Figure 12. Long-run average reward rate by M ($\delta = 10, \mu = 1/30$) 30
Figure 13. Transition diagram when the system is in “Low” or “Mid” risk transient states 38
Figure 14. Transition diagram when the system is in “High” risk transient states 38
Figure 15. Transition diagram when the system is in “Travel” transient states 39
Figure 16. The optimization algorithm 43
Figure 17. The solution for λ_{33} in round 1 46
Figure 18. The changes of the rate λ_{ij} during 10 rounds 47
Figure 19. The changes of the objective value: *P(1,1,1,1)* during 10 rounds 47

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

Table 1. Combinations of different forms of the effective firing rate functions......23

Table 2. Status of each component and the number expression 34

Table 3. The weights and win probabilities in (*, 3, *, 3) states when $\lambda = 0$ 46

Table 4. The updated vector λ^* after round “1” 46

Table 5. The converged optimal vector λ^* 47

Table 6. A result with increasing $\mu_{HR}(j)$ 49

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

BRL	bomb release line
CDF	cumulative distribution function
CTMC	continuous-time Markov chain
DPRK	Democratic People's Republic of Korea (North Korea)
DTMC	discrete-time Markov chain
IID	independent and identically distributed
MOE	measure of effectiveness
PDF	probability density function
ROK	Republic of Korea (South Korea)

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Today, artillery weapons are self-propelled guns that can be easily moved during battle. This leads to the "shoot-and-scoot" tactic: a battery (Blue) fires a small number of rounds at the enemy (Red) and then Blue moves to avoid counter-fire. The shoot-and-scoot tactic is an important maneuver, but there appears to be limited quantitative analysis on how long Blue should stay in one location before moving to a new location. Currently, commanders use their experience and intuition to determine when the artillery should change locations. Most commanders are risk averse, so they tend to move frequently to avoid the enemy's counter-fire. Unfortunately, firing a small number of rounds of artillery and moving rapidly to another position has several drawbacks: a low rate of fire at the enemy and limited opportunities to improve accuracy by adjusting the aim to previous rounds. The benefit of quickly moving positions is a lower risk that Red will detect the location of Blue's artillery. Since tradeoffs exist, it is difficult to determine how long the artillery should remain in one position before moving in order to maximize the benefits and minimize the risk. This thesis focuses on the cost-benefit tension of firing from the same spot over a prolonged period of time. We formulate a model that examines when an artillery force should move positions.

For concreteness, we focus on a particular situation where Blue artillery initially fires on Red. We assume that Blue has the ability to move quickly to another position and Blue has many available positions; for simplicity, we assume Blue never needs to revisit a previous position. In addition, Blue has some information about Red's initial location and capabilities, such as the number of Red's weapons and Red's power. Unlike Blue, Red is stationary (i.e., Red cannot move to another position). Red keeps its position until it is destroyed. Red has sensors (e.g., radars) that can detect the origin of Blue's shells. This allows Red to eventually launch counter-fire at Blue, and Red's accuracy will improve if Blue stays in the same location. To formulate this particular problem, we utilize stochastic analysis as our main approach since there are many uncertainties. Artillery performs "area fire" or "indirect fire" so the probability of a direct hit on the target is very low to start with. Thus, Blue improves the artillery's accuracy by adjusting

the weapons to improve the hit probability. Even if Blue has perfect aim after a number of rounds, however, uncertainty with environmental conditions (e.g., wind velocity and direction) impacts the ammunition. In addition, there is uncertainty with how well Red's counterbattery radars detect the fire and how quickly Red can return fire. We utilize continuous-time Markov chains (CTMC) to analyze this artillery engagement.

In this thesis, we formulate two models to analyze shoot-and-scoot policies for artillery forces. A primary component of our models is "risk," which increases over time when Blue stays in the same position. The risk represents Red's effective firing rate, which is the rate that Red fires rounds multiplied by the probability a round hits Blue. Over time both the gross rate and hit probability may increase as Red homes in on Blue's location. Our first model assumes the battle evolves over a long period of time and defines states according to Blue's risk level. Blue initially fires in a low risk state. Gradually, the risk increases to medium and then high if Blue does not move. When Blue moves to a new location, the risk level resets back to the lowest state. During the transit to a new location Blue faces no risk from Red fire, but poses no threat to Red because Blue does not fire while moving. To determine the optimal move policy, we examine several different objective functions that consider both risk and firing rate. The main objective of this model is to limit Blue's exposure to a higher risk.

Our second model focuses on the probability that Blue will win the battle during a limited time-window scenario. In this model, we incorporate the health of both Blue and Red. If one side's health level decreases to the lowest level, then that side retreats. We also impose a finite battle length. The battle does not go on for an arbitrarily long time: Blue must force Red to retreat within a finite time window. The objective of this model is to maximize the probability Blue wins (i.e., forces Red to retreat). The decision variables in both models are the rates at which Blue moves. In the win-probability model, Blue has more decision variables, as Blue can tailor its move decision based on Blue's health and the stage of the battle.

We explore these two models numerically using realistic parameter values. We fix the expected time from the lowest risk level to the highest risk level at 30 minutes and we set the expected time to change positions to 10 minutes. In our long-run risk model the

results recommend that Blue should move roughly every 15 minutes on average. In the win-probability model, Blue should move frequently during the early stages of the battle. On the other hand, when Blue's health is high, Blue should remain in the same position.

For realistic parameter values, the general result is that in most situations Blue should spend a reasonable amount of time firing multiple shots in the same location. When we account for time and health, this result becomes even more pronounced. Blue should never move in certain states (e.g., high Blue health, later in the battle). Moving frequently reduces risk to Blue, but limits Blue's ability to inflict damage on Red. Based on these results, we conclude that when artillery forces utilize shoot-and-scoot tactics, they should not move frequently because that decreases their opportunity to improve the accuracy and the probability to win the battle. This result may run counter to the approach of some commanders, who believe they should move frequently to survive and win the battle. These results should provide the commanders with insight about shoot-and-scoot tactics. In addition, our models are straightforward and easy to implement. Therefore, artillery commanders can use our models with real battle data and force capabilities.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

Above all, I really appreciate my fellows who are doing their duty in the military against North Korea while I complete my studies. Without their sacrifices, I could not have done my course successfully.

Two years ago, I could not have expected this awesome experience in Monterey, CA. Although I did not apply for the Naval Postgraduate School master's degree course, I would like to thank the officer who selected me to allow me to have the opportunity to study here. Thanks to his mistake (it might not be), I have completed a very tough matrix of 71 credits (56 grad level credits) in only one year. I would like to thank the Republic of Korea Army and Korea National Defense University for allowing me to have a great time in the United States.

I would like to acknowledge everyone who played a role in my academic accomplishment. First of all, I would like to thank my thesis advisor, Dr. Michael Atkinson, for his guidance, encouragement, and patience over the last year. He made my work for this thesis a lot easier than I thought it was going to be. He consistently allowed this paper to be my own work, but steered me in the right the direction whenever he thought I needed it. Without him, I would have never successfully finished my course.

I would also like to acknowledge Dr. Dashi Singham as the second reader of this thesis, and I am grateful for her very valuable comments.

Finally, I must express my very profound gratitude to my lovely wife, Sunhwa Kim, and adorable children, Junhoo and Woojin, for providing me with unfailing support and continuous encouragement throughout my study. Without their patience and support, I would not have been able to devote the time needed to complete my thesis.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

Joseph Stalin, leader of the Soviet Union from 1922–1952, said that “Artillery is the god of war” during World War II (Holmes et al., 2001). Artillery remains an important component of modern warfare. According to Gautam (2010), over the last decade it has been instrumental in conflicts in Iraq, Afghanistan and elsewhere. This thesis examines a specific type of artillery tactic that has become more prevalent as artillery weapon systems improve their capabilities of fire and counter-fire.

Traditionally, a battery (Blue) fires from a fixed position. The benefits from firing in the same location are improved accuracy and a constant and relatively high firing rate. On the other hand, the risk is that the enemy (Red) may eventually determine where Blue is firing from and take countermeasures to attempt to eliminate the Blue artillery. This thesis focuses on the cost-benefit tension of firing from the same spot over a prolonged period of time.

B. MOTIVATION

We formulate a model that examines when an artillery force should move positions. In the past, moving positions was a relatively low priority compared to quickly and accurately firing on the enemy. When artillery consisted primarily of towed cannons, physically moving artillery equipment from one position to another consumed much time, effort, and manpower. In many cases in the past, it was difficult to detect the origin of an artillery round because an observer had to see evidence of the round in real time. In recent years, however, the maneuverability aspect of artillery has become a crucial aspect of artillery battles. With the advent of modern counterbattery radar systems, the origin of artillery fire can be determined safely and quickly away from the shells. Consequently, if artillery remains in its original position for a long time, it will eventually be hit by the enemy’s counter-fire. Also, moving the artillery to another position is much easier than in previous years. Today, artillery weapons are self-propelled guns so troops can easily move quickly to another location. Consequently, this leads to the “shoot-and-scoot”

tactic, which is explained by Koba (1996) that immediately after firing at a target, the artillery changes location to avoid counter-battery fire (p. 16).

In this thesis, we generalize the shoot-and-scoot tactic so the artillery does not need to move immediately after firing one round. There are benefits to firing multiple rounds from the same position. Since tradeoffs exist when a battery stays in the same position for a short time, it is difficult to decide how long the artillery should remain in one position before moving in order to maximize the benefits and minimize the risk. Firing a small number of rounds of artillery and moving rapidly to another position has several drawbacks: a low rate of fire at the enemy and potentially no chance to improve accuracy by adjusting the aim to previous rounds. On the other hand, moving quickly to another position lowers the risk that the enemy will detect the location of the artillery. The problem is that artillery commanders rely mainly on their experience and intuition in deciding when to move to another position. Moreover, in general, most commanders are risk averse and often will move quickly because they prefer to keep their force safe rather than destroy the enemy. For example, a commander may decide to move immediately after the first shot or as soon as the enemy starts to fire even if the enemy's aiming is inaccurate to start. Frequent moving generates low risk, but it consumes much time and effort and imposes a cost of lost firing with improved accuracy. Therefore, in this thesis, we propose a model that examines how long the commander should fire in the same position before moving in order to maximize the fire rate and its accuracy, and minimize the risk. This quantitative analysis provides the artillery commanders with insight about when they should move.

C. SCOPE

For concreteness, we focus on a particular situation. One artillery force (Blue) initially fires at the enemy (Red). We assume that Blue has an ability to move quickly to another position and Blue has many available positions; for simplicity, we assume Blue never needs to revisit a previous position. In addition, Blue has some information about Red's initial location and abilities such as the number of weapons and their power. Unlike Blue, Red is stationary (i.e., Red cannot move to another position). Red keeps its

position unless it is destroyed. Red has radars that can track Blue's shells and detect their origin. This allows Red to eventually launch counter-fire at Blue.

A real-life situation consistent with these assumptions could be artillery engagement between the Republic of Korea (ROK) and Democratic People's Republic of Korea (DPRK). The ROK has a well-developed self-propelled artillery weapon called the "K9 Thunder." Its max speed is 67km/h and its firing range is 40km ("K9 Thunder Self-Propelled Howitzer," 2014). Although the DPRK has many artillery forces on its forward line, its weapons are old with poor maneuverability. Therefore, the DPRK cannot "scoot" after it "shoots." The DPRK has counter-battery radars that can detect the origin of ROK artillery, however. In this circumstance, the ROK takes the shoot-and-scoot tactic to avoid the counter-fire from the DPRK's artillery.

To formulate this particular problem, we utilize stochastic analysis as our main approach since there are many uncertainties. Artillery performs an "area fire" or "indirect fire" so the probability of a direct hit on the target is very low to start with. Thus, Blue has to improve the artillery's accuracy by adjusting the weapons to improve the hit probability. Even if Blue has a perfect aim after a number of shots, however, uncertainty with environmental conditions (e.g., wind velocity and direction) impacts the ammunition. In addition, there is uncertainty with how well Red's counterbattery radars detect the fire and how quickly Red can return fire. In general, the probability that Red detects the firing location depends on several factors. The U.S. Marine Corps (2002) believes these are target type, range, elevation and number of projectiles being simultaneously tracked (U.S. Marine Corps, 2002). Our primary stochastic machinery is the continuous-time Markov chain (CTMC), which is a very useful and powerful approach to analyze stochastic phenomena.

D. LITERATURE REVIEW

Washburn (2002) presents a general treatise in the area of "Firing Theory." It primarily focuses on computing the kill probability obtained from several shots. It accounts for dispersion and bias errors, which can create dependencies across multiple shots. Under some assumptions, the distribution of final shot locations follows a bivariate

normal density. Washburn allows for feedback, which increases accuracy over time and produces higher hit probabilities. This work is much more detailed than we need. We essentially take these assumptions and analysis for granted and assume that when Blue or Red fire, there is some probability of a hit, and we also allow for those probabilities to change over time in various ways as accuracies improve.

Christy (1969) analyzes small unit infantry combat engagements by developing a firefight model using Lanchester's square law to examine different tactical fire and maneuver policies. His objective is similar to ours at a high level. Christy uses infantry forces and considers a maneuver policy based on "distance" to rush. In our analysis, we use "time" as our basis to move. In addition, Christy's model is deterministic and our model is stochastic.

Sweat (1971) considers a single-shot duel between Blue and Red where they have kill and detection probabilities that vary depending on the distance, which are functions of time. Our probabilities also depend upon time, although we allow for multiple hits in our model. Ravid (1989) studies two alternative modes of defense against attacking aircraft: engagement with a lower kill probability before bomb release line (BRL) versus engagement with a higher kill probability after BRL (1989). Sweat and Ravid take "time" as an important decision point to determine when to respond and have tradeoffs about moving earlier versus later. We consider a similar time-dependent tradeoff.

Kress (1991) proposes a model of a two-on-one duel: two Blue units and one Red unit. Similar to our assumptions, Blue can move but Red is stationary. This work, however, focuses on Red's decision about which Blue to engage and Blue's decision when to move toward Red in order to win the battle. Kress does not consider detection systems such as counter-battery radar; Kress assumes Blue and Red know the other's location.

Duke (1996) presents a discrete-time Markov chain (DTMC) to analyze the effectiveness of a new artillery weapon system, at that time called *Crusader*. Duke focuses on the lifetime of *Crusader*, where it waits for fire missions, executes

survivability moves, conducts resupply and executes fire missions until the *Crusader* is killed. It takes a discrete time approach, whereas our model is continuous time.

Harari (2008) considers the opposite side of this thesis. The scenario is that insurgents attack the defender using mortars and short range rockets and the insurgents use the shoot-and-scoot tactic. The Defender has sensors to detect the insurgents, but the sensors are imperfect. The Defender also has missiles to counter-fire at the insurgents. The defender's tradeoff is launching his missile earlier with less accuracy (and potentially causing collateral damage and wasting a missile) or launching it after some aiming process with more accuracy and risk of being too late because the insurgents have already moved locations. In this situation, Harari presents an analytical probability model and simulation result to support the defender's decision making and suggests a new counter-mortar/rocket tactic. The new tactic is that the defender launches his missile immediately after obtaining an initial rough estimate of the launcher's location from the sensor. To achieve it, the missile should be a "smart weapon" that can update its target location information while in flight. In this thesis, we do not model the specifics of Red's firing tactics in great detail. We assume after a random time Red determines Blue's location and starts to return fire with a hit probability that increases over time.

Park (2015) analyzes artillery tactics that consider the distance from the artillery to a moving target. Park utilizes a Markov model and computes the expected time until a retreat condition is satisfied. Park uses a DTMC, whereas we formulate a continuous-time model. We introduce decision variables related to how frequently Blue should move; whereas Park takes a more descriptive approach to examine the impact the distance has on the retreat condition.

E. THESIS OUTLINE

Chapter II introduces the Long-Run Risk Model, which is a CTMC model for analyzing Blue's moving policy. It focuses on a relatively "long" engagement, determines the move policy, and illustrates numerically with examples. In addition, we also consider a renewal process approach to allow for more realistic assumptions. Chapter III develops another CTMC, the Win-Probability model, which incorporates

more realistic aspects: a limited engagement window and the health status of both Blue and Red. Chapter IV compares and analyzes these two CTMC models and concludes the thesis with discussion of suggested tactics and future works.

II. LONG-RUN RISK MODEL

This chapter formulates a model to analyze when Blue should move its artillery to another position. On one hand, Blue wants to rarely move as it is inefficient and decreases Blue's overall firing rate. On the other hand, Blue should move relatively frequently to avoid high risk circumstances where Red has determined Blue's position with reasonable accuracy. In this chapter, we consider only "Risk" as the factor that increases in time as Blue fires from the same position. We define the risk to Blue as the effective firing rate of Red, which is the rate that Red fires rounds multiplied by the probability a Red round hits Blue. These two quantities (especially the hit probability) will increase in time as Blue stays at the same location. The risk to Red may also increase in time as Blue increases its effective firing rate. However, for most of the analysis in this Chapter, we assume a constant risk to Red (i.e., Blue's effective firing rate is constant). We take a long-run approach to the problem. In the next chapter, we incorporate additional components such as the health of Blue and Red and a limited time horizon into a related, but separate, model. We take a CTMC approach to this problem, and thus we give a brief overview of CTMCs in Section A before describing the model in Section B. After the model description, we analyze the long-run behavior of the model in Section C and extend it to a more general case in Section D. We present results for the CTMC models in Sections E and F. Finally, in Section G, we present a Renewal Process approach to the problem that allows us to relax some of the non-realistic assumptions in the CTMC model.

A. CTMC REVIEW

Consider a discrete-time stochastic process $\{X_n, n = 0,1,2,...,\}$ with X_n taking on values in the *state space*. For concreteness, assume here that the state space is the non-negative integers. If $X_n = i$, then the process is said to be in state *i* at time *n*. Whenever the process is in state *i* at time *n*, there is a fixed probability P_{ij} that it will next be in state *j* at time *n+1*. Knowing the state of the process in previous periods does not convey any additional information about the state in period *n+1*.

Mathematically we have

$$
\begin{aligned}
&P\{\underbrace{X_{n+1}=j}_{future} \mid \underbrace{X_n=i}_{present}, \underbrace{X_{n-1}=i_{n-1},\dots,X_1=i_1,X_0=i_0}_{past}\} \\
&= P\{\underbrace{X_{n+1}=j}_{future} \mid \underbrace{X_n=i}_{present}\} = P_{ij}
\end{aligned}
$$

for all states $i_0, i_1, \dots, i_{n-1}, i, j$ and all $n \geq 0$. Such a stochastic process is known as a discrete-time Markov chain (DTMC). If we know the present state X_n, the future state X_{n+1} is independent of the successive past states $\{X_{n-1}, \dots, X_1, X_0\}$.

There is an analogous stochastic process in continuous time $\{X(t), t \geq 0\}$ called a continuous-time Markov chain (CTMC). The Markov condition for CTMC becomes

$$
\begin{aligned}
&P\{\underbrace{X(t+s)=j}_{future} \mid \underbrace{X(s)=i}_{present}, \underbrace{X(u)=x(u)}_{past}, 0 \leq u < s\} \\
&= P\{\underbrace{X(t+s)=j}_{future} \mid \underbrace{X(s)=i}_{present}\}
\end{aligned}
$$

for all $s, t \geq 0$ and all states $i, j, x(u), 0 \leq u < s$. Again, the future state $X(t+s)$ depends on the process history only through the present state $X(s)$. Ross (2014) introduces several properties of a CTMC as follows.

1. The amount of time the system spends in state *i* before transitioning into a different state is exponentially distributed with rate μ_i.
2. When the process leaves state i, it next enters state j with some probability P_{ij} which must satisfy $P_{ii} = 0$ and $\sum_j P_{ij} = 1$ for all i.
3. The time until the process transitions from state *i* to state *j* has an exponential distribution with rate $q_{ij} = P_{ij} \times \mu_i$ for $j \neq i$. The matrix Q containing the q_{ij} rates is called the infinitesimal generator matrix.

A CTMC can be viewed as a stochastic process that moves from state to state in accordance with a DTMC, but the amount of time it spends in each state, before proceeding to the next state, has an exponential distribution. See chapters 4 and 6 of Ross (2014) for details.

B. MODEL DESCRIPTION

Blue fires at Red for some amount of time and then moves to a new location. After Blue moves to the new location, Red eventually returns fires. As Blue fires in time, Red obtains information through sensors (e.g., radars) or human resources like reconnaissance units about the exact position of Blue. Thus the probability that Red hits Blue with a round increases in time. We define the risk to Blue (henceforth just risk) as Red's effective firing rate: the overall rate Red fires rounds multiplied by the hit probability. The risk increases if Blue stays in the same location, even if Blue never fires, because Red may have reconnaissance units or surveillance assets (e.g., satellites or UAVs) that can pinpoint Blue's location. In our model risk explicitly increases with time. Implicitly we assume this occurs primarily as Red reacts to Blue's artillery fire. However, it may also increase in time for other reasons as mentioned previously. A more accurate (and perhaps complex) model would more directly tie Blue's fire to increases in risk.

After moving to a new location, the risk resets to the lowest level as Blue begins firing from the new position. This follows because we assume Red has limited information about the new location of Blue and thus poses little threat to Blue at this initial firing time. We model this as a CTMC. At any time *t*, the system is in one of the following four states:

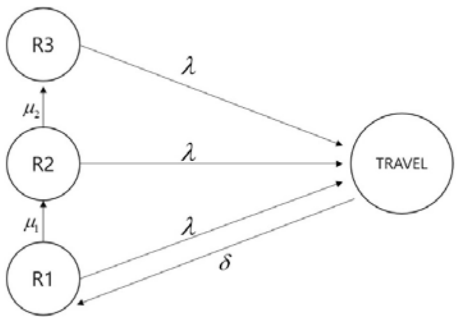
- R1: Low risk (this occurs immediately after traveling to a new position)
- R2: Medium risk
- R3: High risk
- TRAVEL: Blue moves to another position

To reiterate: for each risk state, Red's effective firing rate is constant. In time as Red collects more intelligence about Blue's location, the effective firing rate increases (e.g., hit probability increases), and hence the system transitions to the next higher risk level. For most of this chapter, we do not explicitly specify the effective firing rate of Red or Blue. We assume Blue has a constant firing rate and Red has a firing rate that increases with the risk level. We only consider the risk level and assume Blue prefers to be in lower risk levels. In sections E.4 and F.4 we analyze a scenario with specific effective firing rates for both Red and Blue. Here, we only consider three risk levels. In the next section, we generalize to an arbitrary number of risks levels. Blue has a lower probability to be hit by Red in the low risk state than in medium or high risk states.

We assume the system starts at time 0 when Blue arrives to a new position and starts firing. This corresponds to the low risk level R1. Gradually, risk increases to Blue as Red better determines Blue's position. As we model this as a CTMC, we assume the times until the risk increases by one level are exponentially distributed with μ_i, where index i represents the current risk level. The time between risk level increases corresponds to the time it takes Red to improve its effective firing rate, which involves re-aiming to increase accuracy, interpreting the radar signals, processing surveillance information, and switching modes to fire at a faster rate. We assume that the time until Blue moves (and enters the travel state) is also exponential with rate λ. The key decision for Blue is setting λ, which dictates Blue's move policy. Finally, the travel time is also exponential. It may be unrealistic to model all times as exponential, especially the movement and travel times. We want to formulate an analytically tractable approach to the problem, however. We discuss non-exponential times at the end of this chapter, which may provide more realistic settings.

Figure 1 shows the possible transitions between these states. Whenever in the low or medium risk level (state R1 or R2), Blue transitions to TRAVEL with rate λ and the next higher risk state with rate $\mu_i (i = 1, 2)$. Whenever in the highest risk level (state R3), Blue only transitions next to the TRAVEL state with rate λ. The move rate λ does not depend upon the risk level; we discuss this assumption in more detail at the end of this

section. Whenever in state TRAVEL, Blue only transitions to R1 state with rate δ. This assumes that Blue moves to a totally new position whenever Blue completes its travel. In other words, once Blue arrives back to the R1 state, the system will be considered as new, regardless of what occurred previously. As the system evolves in time, all times are independent and have an exponential distribution with the rates presented earlier.



$$Q = \begin{array}{c} \\ R1 \\ R2 \\ R3 \\ T \end{array} \begin{array}{c} \begin{array}{cccc} R1 & R2 & R3 & TRAVEL \end{array} \\ \begin{pmatrix} -(\mu_1+\lambda) & \mu_1 & 0 & \lambda \\ 0 & -(\mu_2+\lambda) & \mu_2 & \lambda \\ 0 & 0 & -\lambda & \lambda \\ \delta & 0 & 0 & -\delta \end{pmatrix} \end{array}$$

Figure 1. Transition diagram and its infinitesimal generator matrix Q

The travel rate δ and the risk rates μ_i depend on the situation on the battlefield and we take them to be exogenous parameters. In particular, the risk rates $\mu_i (i = 1, 2)$ may differ: μ_1 may be greater than μ_2 and vice versa. For example, imagine that finding Blue's position is more difficult for Red than improving Red's firing accuracy once Blue has been located. In that case, μ_1 will be small and Blue will stay in the lowest risk state for a while (probabilistically) as it takes time for Red to locate Blue. The risk level will transition very quickly (probabilistically) from Medium to High because μ_2 will be much larger, however. In other cases, it may be difficult to improve the accuracy after reaching some point of accuracy, which would correspond to a large μ_1 and a small μ_2. Later in this chapter, we present results of several experiments varying μ_i.

While δ and μ_i are fixed inputs, the parameter λ is a decision variable of the Blue commander. We make an important assumption that λ is constant across all risk levels. That is, the commander cannot tailor his move decision based on the current risk. There are a few possible justifications for this assumption. For operational reasons the

move decision cannot be made in real time but must be set ahead of time (e.g., as soon as the arrival to new location). Another reason is the Blue commander may not actually know the true risk level in real time. If the commander knew the true risk level, he would probably move more frequently in higher risk states (i.e., higher λ in R3) and less frequently in lower risk states (i.e., lower λ in R1). The risk level is Red's effective firing rate. While Blue can observe the accuracy of current incoming fire, we feel that does not provide enough information to make an informed estimate of Red's effective firing rate, and hence the current risk level, in real time. Therefore, we assume that since Blue does not know the true risk level, it can only make one move decision and hence one λ parameter.

C. THE LONG-RUN BEHAVIOR

In order to compute the optimal move policy (i.e., the optimal λ), we need to specify an objective. We assume that this battle goes on for an infinite amount of time, or at least long enough such that the infinite time horizon is reasonable. One possible objective function is the proportion of time the Blue artillery is in the low risk state. In order to compute an objective function about the long-run behavior of the system, we first need to compute the limiting distribution of the CTMC. We denote the long-run proportion of time the system is in state *i* as π_i. To compute the π_i, we solve the balance equations of the CTMC. Roughly speaking these balance equations specify that the rate at which a CTMC transitions out of a state must equal the rate at which the CTMC transitions into the state. See section 6.5 of Ross (2014) for more information on solving for the limiting distribution. For instance, in the TRAVEL state the incoming rates are $(\pi_{R1}+\pi_{R2}+\pi_{R3})\lambda$ and the outgoing rate is $\pi_{TRAVEL}\delta$. Since $\pi_{R1}+\pi_{R2}+\pi_{R3}+\pi_{TRAVEL}=1$, it yields $\pi_{TRAVEL}\delta=(1-\pi_{TRAVEL})\lambda$ and then produces

$$\pi_{TRAVEL}=\frac{\lambda}{\lambda+\delta}$$

The same approach in every state yields

$$\pi_{R1}(\lambda+\mu_1)=\pi_{TRAVEL}\delta \qquad \pi_{R1}=\frac{\delta}{\lambda+\mu_1}\pi_{TRAVEL}=\frac{\delta}{\lambda+\mu_1}\times\frac{\lambda}{\lambda+\delta}$$

$$=\frac{\lambda}{\lambda+\mu_1}\times\frac{\delta}{\lambda+\delta}$$

$$\pi_{R2}(\lambda+\mu_2)=\pi_{R1}\mu_1 \quad \rightarrow \quad \pi_{R2}=\frac{\mu_1}{\lambda+\mu_2}\pi_{R1}=\frac{\mu_1}{\lambda+\mu_2}\times\frac{\delta}{\lambda+\mu_1}\times\frac{\lambda}{\lambda+\delta}$$

$$=\frac{\lambda}{\lambda+\mu_2}\times\frac{\mu_1}{\lambda+\mu_1}\times\frac{\delta}{\lambda+\delta}$$

$$\pi_{R3}\lambda=\pi_{R2}\mu_2 \qquad \pi_{R3}=\frac{\mu_2}{\lambda}\pi_{R2}=\frac{\mu_2}{\lambda}\times\frac{\mu_1}{\lambda+\mu_2}\times\frac{\delta}{\lambda+\mu_1}\times\frac{\lambda}{\lambda+\delta}$$

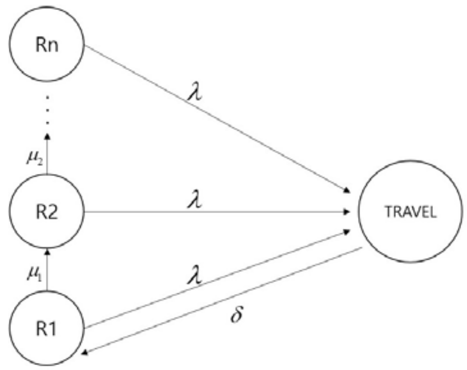
$$=\frac{\mu_2}{\lambda+\mu_2}\times\frac{\mu_1}{\lambda+\mu_1}\times\frac{\delta}{\lambda+\delta}$$

One interpretation for the long-run proportion of time in the lowest risk state R1 (π_{R1}) is the probability Blue is not traveling $(\frac{\delta}{\lambda+\delta})$, multiplied by the probability Blue moves before increasing to state R2 $(\frac{\lambda}{\lambda+\mu_1})$. Similarly, the long-run proportion of time at state R2 (π_{R2}) is the probability Blue is not traveling $(\frac{\delta}{\lambda+\delta})$, multiplied by the probability we reach R2 before moving $(\frac{\mu_1}{\lambda+\mu_1})$, multiplied by the probability Blue moves before increasing the risk to state R3 $(\frac{\lambda}{\lambda+\mu_2})$. A similar interpretation holds for the limiting distribution for R3.

D. GENERAL NUMBER OF RISK LEVELS

In the previous section, we arbitrarily defined three risk states: low, medium, and high. In this section, we increase the number of risk states. There are two extreme risk points: no risk and the highest risk. How many states do we need between them to adequately represent reality? Three may be enough, but perhaps 10 or even 100 or 1,000

would be better. It is possible a more refined risk level granularity may capture the real risk better. From now on, we use *n* as the number of the risk states. Figure 2 shows the transition diagram for this generalized model.



$$
Q = \begin{array}{c} \\ R1 \\ R2 \\ \vdots \\ Rn \\ T \end{array}
\begin{array}{c}
\begin{array}{ccccc} R1 & R2 & & Rn & T \end{array} \\
\begin{pmatrix}
-(\mu_1+\lambda) & \mu_1 & \cdots & 0 & \lambda \\
0 & -(\mu_2+\lambda) & \cdots & 0 & \lambda \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & -\lambda & \lambda \\
\delta & 0 & \cdots & 0 & -\delta
\end{pmatrix}
\end{array}
$$

Figure 2. Advanced CTMC model

Going through similar steps as in the previous section, we can compute the limiting distribution.

$$\pi_{TRAVEL} = \frac{\lambda}{\lambda+\delta} \qquad \pi_{FIRING} = \frac{\delta}{\lambda+\delta}$$

$$\pi_{R1} = \frac{\delta}{\lambda+\mu_1}\pi_{TRAVEL} = \frac{\lambda}{\lambda+\mu_1} \times \frac{\delta}{\lambda+\delta}$$

$$\pi_{Rk} = \frac{\mu_{k-1}}{\lambda+\mu_k}\pi_{Rk-1} = \frac{\lambda}{\lambda+\mu_k} \times \prod_{i=1}^{k-1}\left(\frac{\mu_i}{\lambda+\mu_i}\right) \times \frac{\delta}{\lambda+\delta}, \qquad k = 2,3,\ldots,n-1$$

$$\pi_{Rn} = \frac{\mu_{n-1}}{\lambda}\pi_{Rn-1} = \prod_{i=1}^{n-1}\left(\frac{\mu_i}{\lambda+\mu_i}\right) \times \frac{\delta}{\lambda+\delta}$$

E. OBJECTIVE FUNCTIONS FOR OPTIMIZATION

To determine the optimal move policy λ, we need to define an objective function that adequately captures the tension the commander faces. Moving frequently is inefficient and avoids performing the mission task: firing on the enemy. Moving infrequently exposes Blue to increased risk via increased effective fire from Red, however. In this section, we introduce several possible objective functions to maximize Blue's benefit. We numerically examine these objective functions in Section F.

1. $\underset{\lambda}{Max} \quad Z_1 = \pi_{R1}$

The state R1 is the best state for Blue: Blue fires on Red in a low risk setting. Consequently, we first consider this simple objective function.

2. $\underset{\lambda}{Max} \quad Z_2 = \pi_{R1} - \pi_{Rn} - \pi_{TRAVEL}$

The commander desires state R1 but also wants to avoid the highest risk states. Furthermore, the commander wants to avoid excessive travel because Blue is not firing on Red when Blue travels. This objective is a modification of the first one that penalizes the time traveling and the time in the highest risk state Rn.

3. $\underset{\lambda}{Max} \quad Z_3 = -\pi_{Rn} - \pi_{TRAVEL} + \sum_{i=1}^{n-1} w_i \pi_{Ri}$

For a small number of risk levels (small *n*), objectives 1 and 2 may suffice. For larger *n*, however, those objectives ignore all the intermediate states between R1 and Rn. The lower risk levels may provide benefits and we may want to penalize the higher risk states. In this objective, we use a weight for all risk levels except the highest. We assign higher weights to lower risk states since Blue wants to spend more time in these lower risk states. Also, the sum of all weights is 1 $(\sum_{i=1}^{n-1} w_i = 1)$.

4. $\underset{\lambda}{Max} \quad Z_4 = \sum_{i=1}^{n} (f_{B,Ri} - f_{R,Ri}) \pi_{Ri}$

We have primarily focused on the increased risk to Blue from staying in the same location for a long time. There may be increased benefits to Blue in staying in the same

location for a long time, however. We assume Red's effective firing rate increases over time as Red's accuracy at hitting Blue increases; the same could hold for Blue's accuracy and effective fire rate. In this objective, we consider the actual effective firing rate *f* of Red and Blue corresponding to each risk level. $f_{B,Ri}$ means the firing rate of Blue in state Ri. Moreover, $(f_{B,Ri} - f_{R,Ri})$ is the relative firing rate for Blue. Presumably both $f_{B,Ri}$ and $f_{R,Ri}$ increase with risk level Ri. Therefore, the Blue commander may want to overwhelm Red by staying longer to achieve greater relative firing rates at higher risk states. We will examine different forms of the firing rate function: linear, concave, convex. A concave function may describe the situation where there is a quick learning curve to initially improve accuracy to moderate levels, but it is much more difficult to increase from moderate accuracy to high accuracy. On the other hand, a convex function can model the situation where it is difficult to initially calibrate the artillery, but thereafter the accuracy improves quickly. We consider this further in the next section.

F. NUMERICAL DEMONSTRATION

To implement this model, we use the R: A language and environment for statistical computing (R core team, 2016). We use the `optimize()` function to find the optimal solutions in R. We experiment with different *n* to examine the impact when we have more levels. For the purpose of comparison and analysis, however, we fix the travel rate at δ and assume the expected time to transition from the lowest risk to highest risk when Blue cannot travel (i.e., $\lambda = 0$) is constant for any value of *n*. We denote this expected time between lowest and highest risk states as *T*. If t_i is the amount of time spent in state Ri for $i = 1, 2, ..., n-1$, then

$$T = \sum_{i=1}^{n-1} E(t_i \mid \lambda = 0) = \frac{1}{\mu_1} + \cdots + \frac{1}{\mu_{n-2}} + \frac{1}{\mu_{n-1}}$$

In other words, we split the fixed expected amount of time from lowest risk to highest risk into $n-1$ parts by varying μ_i appropriately with *n*. For example, if $n = 3$

and $T=30$, then $\mu_1=\mu_2=1/15$ is one of the possible values for μ_i. We will use the following travel rate δ and time T as our base case values.

$$\delta=\frac{1}{10}, \qquad T=30$$

There are several possibilities for defining the μ_i: increasing, decreasing or constant. We use the following three definitions for μ_i for the three scenarios. Recall the expected time T is a constant.

- Increasing: $\mu_i=\frac{1}{T}\times\frac{\sum_{j=1}^{n-1} j}{n-i}=\frac{1}{T}\times\frac{n(n-1)}{2(n-i)}, \quad for \quad i=1,2,\ldots,n-1$
- Decreasing: $\mu_i=\frac{1}{T}\times\frac{\sum_{j=1}^{n-1} j}{i}=\frac{1}{T}\times\frac{n(n-1)}{2i}, \quad for \quad i=1,2,\ldots,n-1$
- Constant: $\mu_i=\frac{n-1}{T}, \quad for \quad i=1,2,\ldots,n-1$

For example, if n = 10 and μ_i is increasing, we use $\mu_1=0.167,\ldots,\mu_8=0.75$ and $\mu_9=1.5$. Before proceeding, we discuss the limiting case when $n\to\infty$. A decision-maker may want to incorporate finer resolution in modeling changes in Red's effective firing rate (i.e., risk), and thus we explore the limiting behavior, which may be a reasonable approximation even for modest values of *n*. For all three scenarios (increasing, decreasing, constant) the transition rates $\mu_i\to\infty$ when $n\to\infty$. It turns out that in the limit for all three cases, the time between entry into the lowest risk state and entry to the highest risk state is deterministic. To prove this, we compute the mean and variance for this quantity and show that the variance converges to 0 in the limit.

We focus on the constant case for concreteness, but the proof for the other cases is similar. We define the rate for the time between transitions as $\mu_i=\mu^C=\frac{n-1}{T}$ for all risk

levels. Since all the random variables t_i are exponentially distributed and independent, the random variable *W*, which is the total amount of time from the lowest risk state to the highest, follows a gamma distribution with shape $n-1$ and rate μ^C as follows.

$$W = \sum_{i=1}^{n-1} t_i \sim Gamma(n-1, \mu^C)$$

Consequently, the mean and variance of *W* go to *T* and 0 respectively.

$$\lim_{n\to\infty} E[W] = \lim_{n\to\infty} \frac{n-1}{\mu^C} = T \qquad \lim_{n\to\infty} Var[W] = \lim_{n\to\infty} \frac{n-1}{(\mu^C)^2} = \lim_{n\to\infty} \frac{T^2}{n-1} = 0$$

This implies that if we send $\mu_i \to \infty$, the random variable *W* becomes the deterministic value *T*. We will examine this deterministic case in Section G. Therefore, in the following numerical examples we only look at smaller *n* (3 or 10). We primarily focus on the increasing and decreasing patterns and later discuss the constant case. We now focus on the specific objective functions introduced in Section E.

1. $\underset{\lambda}{Max} \quad Z_1 = \pi_{R1}$

Figure 3 shows several results using different *n* and increasing or decreasing μ_i. We compute the optimal λ numerically, which is a straightforward exercise as it is a one-dimensional optimization problem. When we increase *n*, the optimal objective value (Z_1) (i.e., π_{R1}) decreases and the rate λ increases. This means that Blue moves more frequently because for larger *n*, μ_i is larger and thus Blue remains in the lowest risk state for a (probabilistically) smaller time before transitioning to risk state 2. The only way to return to risk state 1 is by moving and hence λ increases. Unfortunately, larger λ corresponds to a greater long-run proportion of time Blue spends in state TRAVEL (π_{TRAVEL}); when *n* = 10, Blue spends over 0.6 of its time traveling and thus fires at Red less

than 0.4 of the time. This is a somewhat disappointing result as Blue needs to fire at Red to complete the mission.

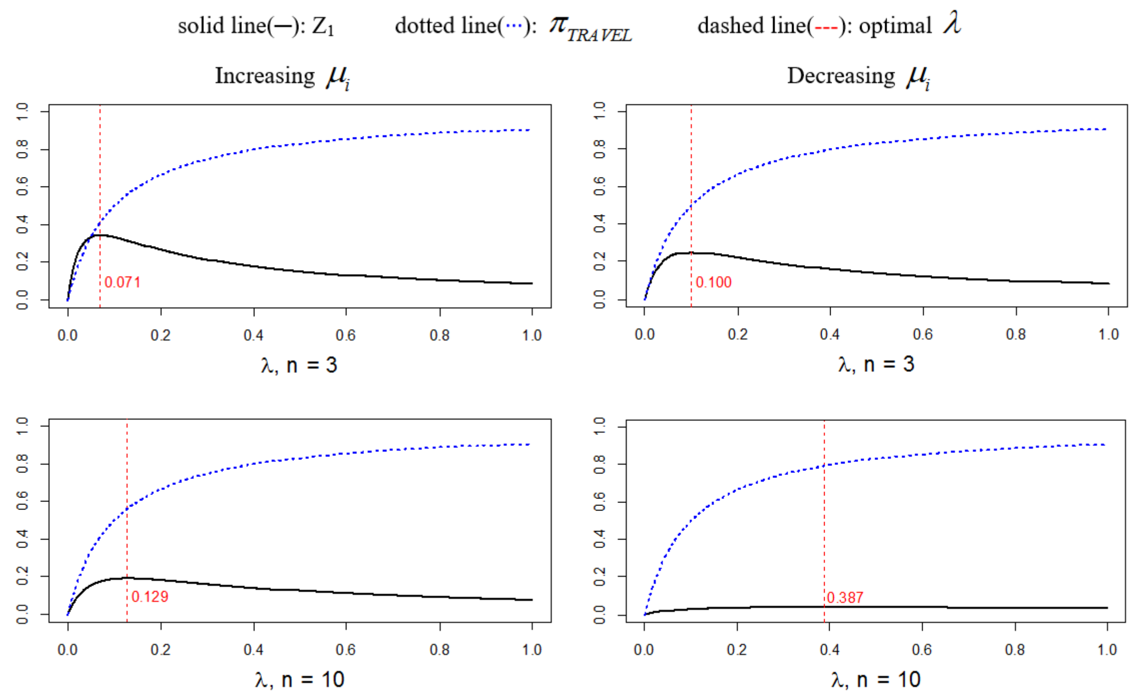


Figure 3. Plots using different n and change patterns of μ_i

For both *n=3* and *n=10*, the optimal rate λ is larger for decreasing μ_i. The reason is that when μ_i decreases, Blue leaves risk level 1 very quickly compared to the increasing case; μ_1 is greater in the decreasing case compared to the increasing case. In order to return to the low risk state, Blue needs to move and that is why λ is greater in the decreasing case. When n = 10, $\lambda^* = 0.129$ with increasing μ_i and $\lambda^* = 0.387$ with decreasing μ_i. This corresponds to Blue moving its position on average 7.8 minutes (1/0.129) after firing commences for the increasing μ_i scenario and 2.6 minutes (1/0.387) for decreasing μ_i.

2. $\underset{\lambda}{Max} \quad Z_2 = \pi_{R1} - \pi_{Rn} - \pi_{TRAVEL}$

Figure 4 shows the results. The optimal λ is much smaller compared with the previous one because this objective penalizes π_{TRAVEL}, which increases with λ. Inspection of Figure 4 reveals an issue with using this objective for larger *n*. As we vary λ from 0 to 1, both π_{TRAVEL} and π_{Rn} range from nearly 0 to nearly 1. For large *n*, however, π_{R1} is close to zero for all values of λ. Thus, for larger values of *n*, the optimization problem simplifies to minimizing $\pi_{Rn} + \pi_{TRAVEL}$, which ignores the time spent firing in low risk (π_{R1}). While it may be a valid objective to only focus on penalizing those two states, it does suggest that for larger *n*, we need to consider more "good" states other than R1. The next subsection considers a different objective function that does just this.

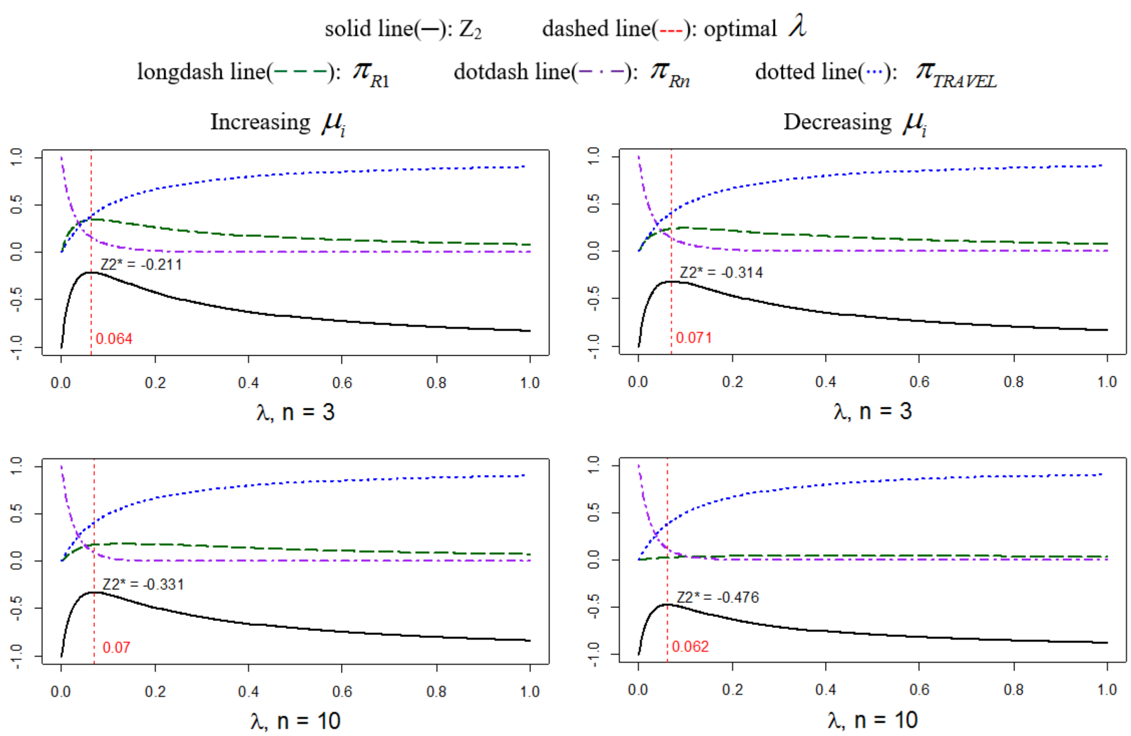


Figure 4. Plots using the objective function Z_2

3. $$\underset{\lambda}{Max} \quad Z_3 = -\pi_{Rn} - \pi_{TRAVEL} + \sum_{i=1}^{n-1} w_i \pi_{Ri}$$

An issue with the first two objective functions is that they only put a premium on being in the lowest risk state. When we only consider three risk states this may be reasonable, but for larger *n*, Blue may consider other lower risk levels (e.g., 2, 3) acceptable. In this objective function, we give different weights at each risk state except the highest risk state, where we penalize it as in the previous objective. We could also assign a negative weight to higher levels, such as $n-1$, $n-2$, etc., but we decide to assign them a (probably small) positive weight. The reason is that, in general, for most parameters of interest, the proportion of time Blue spends in an intermediate risk level right below the highest (risk level $n-1$) is almost zero. It is possible Blue spends a significant amount of time in the highest risk state, however. Figure 5 illustrates the limiting distribution for the various risk levels for different parameters.

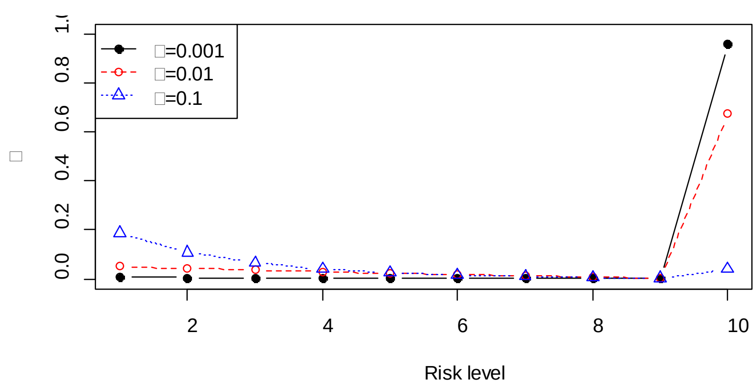


Figure 5. Examples of distribution of π_i, n = 10

For concreteness, we define the weights as follows, which puts higher weights on lower levels.

$$w_i = \frac{n-i}{\sum_{k=1}^{n-1} k} = \frac{2(n-i)}{n(n-1)}, \qquad i = 1, 2, \ldots, n-1$$

For example, when $n=10$, $w_1=0.2$, $w_2=0.178$,… and $w_9=0.02$. Further, the objective function becomes

$$\underset{\lambda}{Max} \quad Z_3 = -\pi_{Rn} - \pi_{TRAVEL} + \frac{2}{n(n-1)}\sum_{i=1}^{n-1}(n-i)\pi_{Ri}$$

Figure 6 shows the results. For this objective, we have a fairly stable optimal solution λ that does not depend on the number of risk states *n* and the pattern of μ_i.

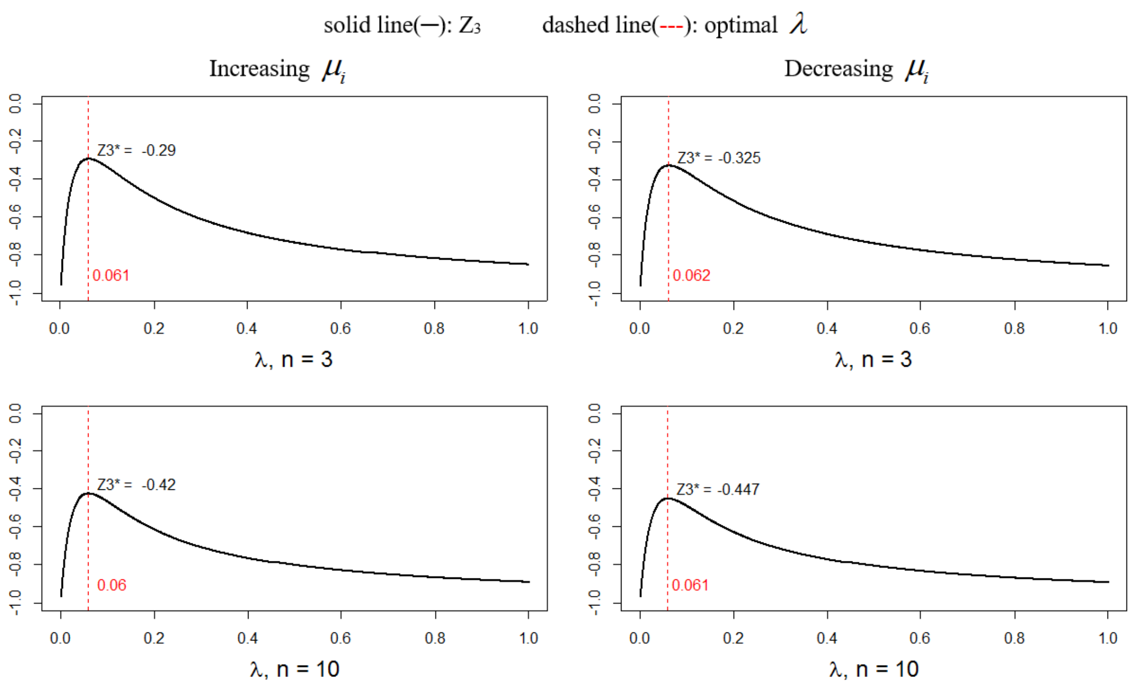


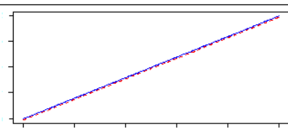
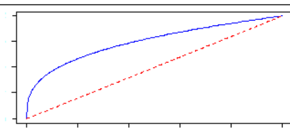
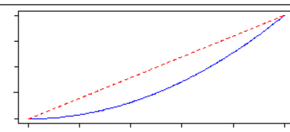
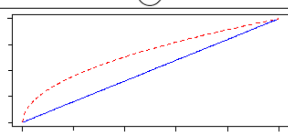
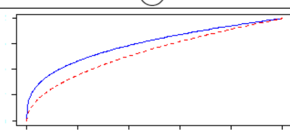
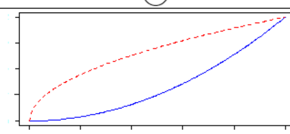
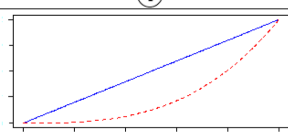
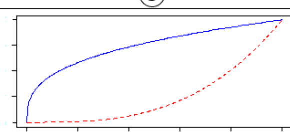
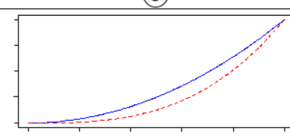
Figure 6. Plots using the objective function Z_3

4. $\underset{\lambda}{Max} \quad Z_4 = \sum_{i=1}^{n}(f_{B,Ri} - f_{R,Ri})\pi_{Ri}$

Both Blue and Red have an effective firing rate *f* that increases with the risk level. We assume the effective firing rate takes one of three forms: concave, convex, linear. This produces nine combinations of the firing rate structure between Blue and Red, as illustrated in Table 1. We assume that the minimum effective firing rate $f_{\min}$ is 2 per minute and the maximum firing rate $f_{\max}$ is 10 per minute on both sides and $n=10$.

Also, we use different specific functions for the concave and convex functions so they do not trivially overlap. The x-axis in Table 1 corresponds to the risk level (1 to 10) and y-axis the firing rate (min: 2 and max: 10). The functional form we use is $f_{\min} + (f_{\max} - f_{\min})(\frac{i-1}{9})^{a}$ where a is 1 (Linear), 1/3 (Concave) and 2 (Convex) for Blue, and 1 (Linear), 1/2 (Concave) and 3 (Convex) for Red.

Table 1. Combinations of different forms of the effective firing rate functions

Red \ Blue	Linear	Concave	Convex
Linear	 ①	 ②	 ③
Concave	 ④	 ⑤	 ⑥
Convex	 ⑦	 ⑧	 ⑨

solid line(—): Blue dashed line(---): Red

We classify the nine situations illustrated in Table 1 according to who has the higher firing rate. We number the 9 figures in Table 1 moving left to right and up to down so we can we refer to specific cases by the label in discussions. First, Blue has the higher effective firing rate (i.e., cases labeled 2, 5, 7, 8 and 9 in Table 1). Second, Red has the higher effective firing rate (i.e., labels 3, 4 and 6 in Table 1). We ignore combination 1 as it generates an objective value of 0 for any λ as the effective firing rates for Blue and Red are the same across all risk levels. We look at the second case first, where Red has a higher effective firing rate. For convenience, we use the constant μ_i scenario.

Figure 7 shows the results when Red has the higher effective firing rate. This situation, as presented in Table 1, produces somewhat trivial results as Blue should never move ($\lambda^* = 0$) and the objective value is zero ($Z_4^* = 0$). The best Blue can do in these cases is have the same effective firing rate as Red, which occurs in either the lowest risk state or the highest risk state (see Table 1). For all other risk states, Red's effective firing rate dominates Blue's effective firing rate. Consequently Blue has two choices: move very frequently to only fire in the low risk level or never move so the situation remains in the high risk level.

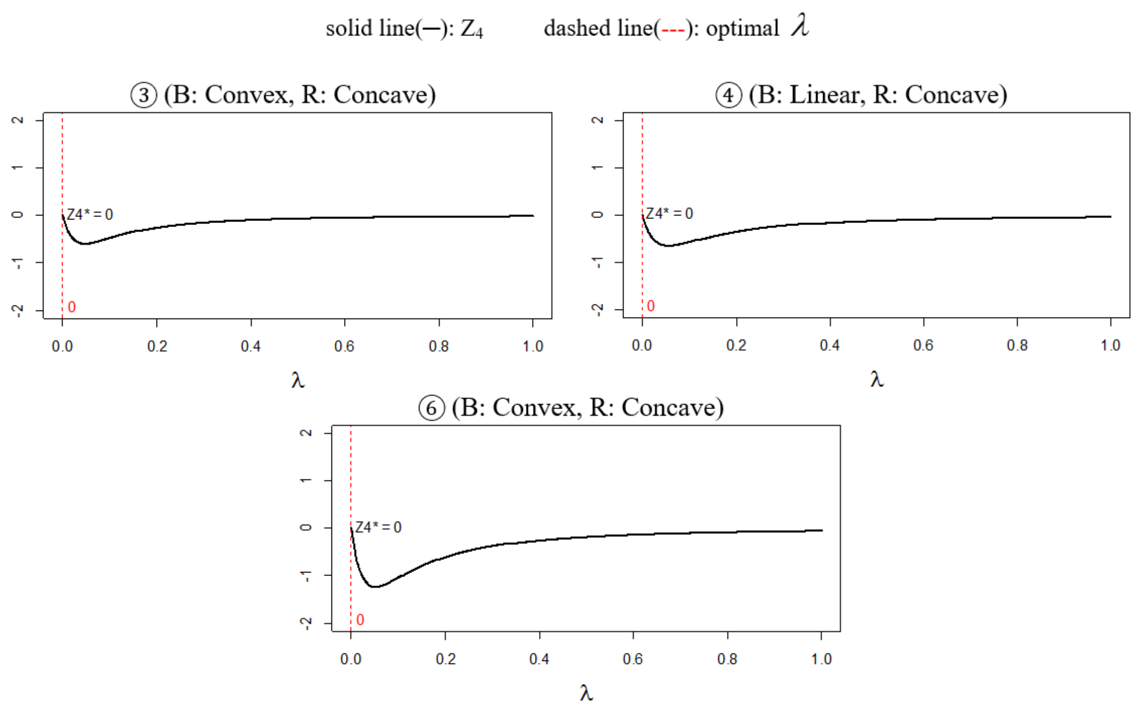


Figure 7. Plots when Red has a higher effective firing rate

Figure 8 shows the result when Blue has a higher effective firing rate. In this case, the optimal solutions are not the same. The optimal λ lie in (0.043, 0.079) across the different scenarios. In combination 8, the optimal objective value is the highest, followed by cases displayed in blocks 2, 7, 5 and 9 in order. Intuitively, Blue wants to spend a large fraction of time in states where Blue's effective firing rate is much higher than Red's rate because that is the objective of interest. Figure 9 shows the relative effective

firing rates ($f_{B,Ri} - f_{R,Ri}$) by the risk states. In combination 5, the relative effective firing rate is higher in the early risk levels and the highest at the state R2. This explains why combination 5 has highest optimal λ in Figure 8. On the other hand, Blue prefers the states near state R7 in combination 9, which leads to the lowest optimal λ among the considered scenarios.

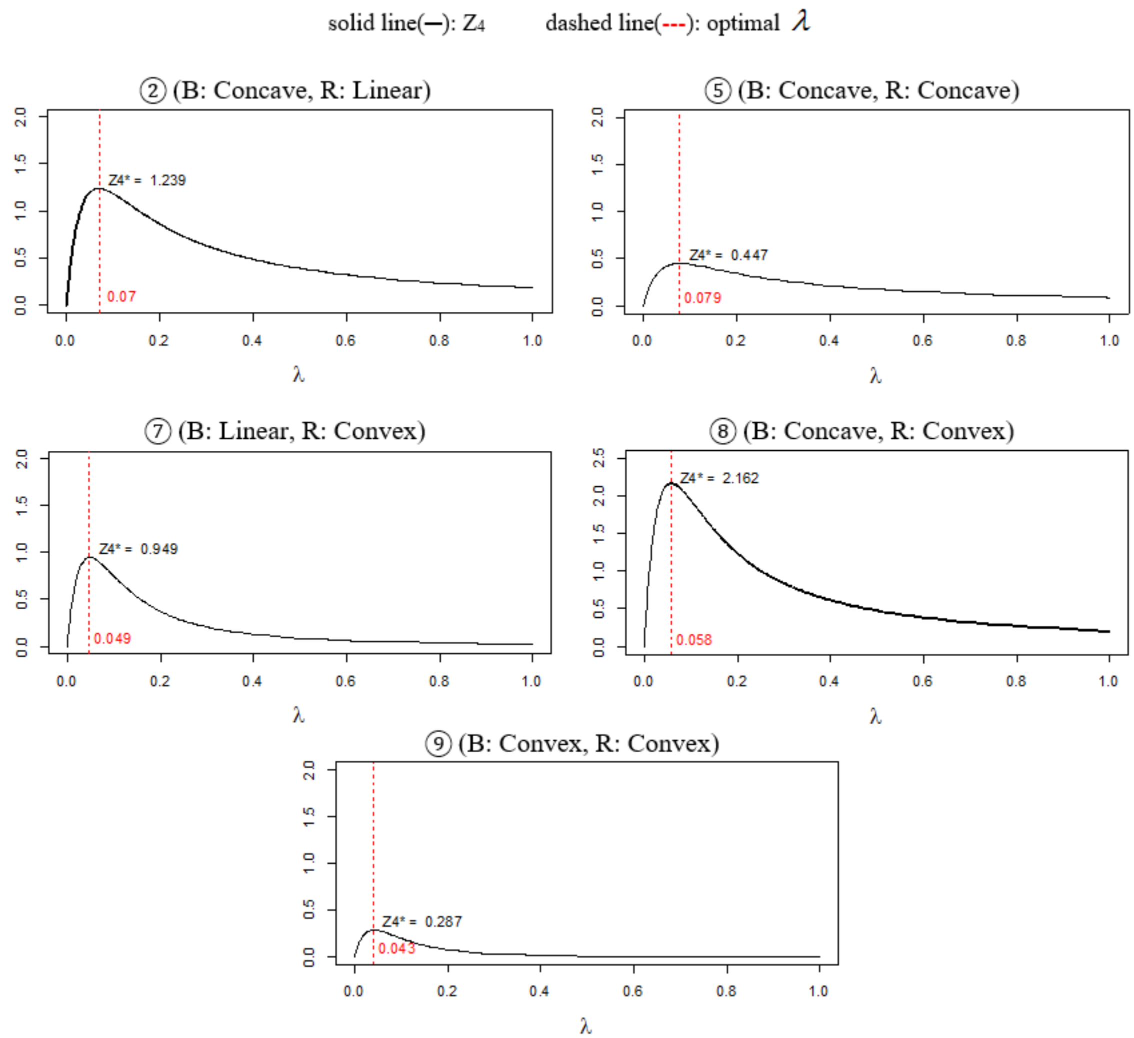
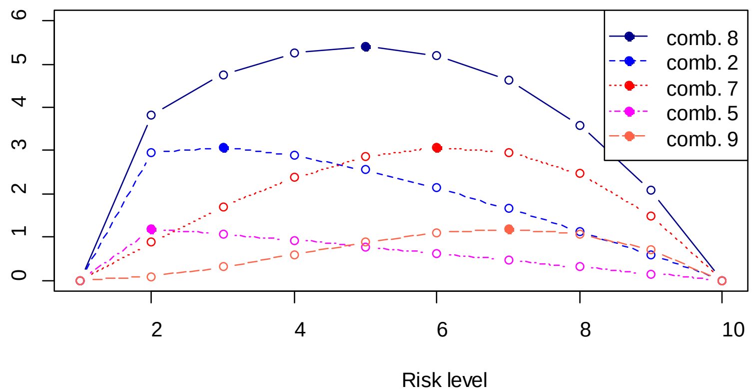


Figure 8. Plots when Blue has a higher effective firing rate



The legend labels correspond to the number labels in Table 1.

Figure 9. Relative effective firing rate ($f_{B,Ri} - f_{R,Ri}$) in the cases when Blue has a higher firing rate

G. REWARD RENEWAL PROCESS

In previous sections, we assume all times are exponentially distributed. This is not realistic, so we allow for some non-exponential times by using a renewal approach. A renewal process is a counting process $\{N(t), t \geq 0\}$, and X_n denotes the interarrival time between ($n-1$)st and nth events. The interarrival times must be nonnegative independent and identically distributed (IID) random variables. The key in constructing a renewal process is defining a renewal point where the process restarts or regenerates. In our case this happens whenever Blue moves and enters the low risk state. See chapter 7 of Ross (2014) for details.

1. Deterministic T

First, we assume the time to transition from the lowest risk to the highest risk is deterministic. As discussed earlier, the CTMC model approaches this for large *n*. With this approach, we define two states: risk (firing) and travel. After traveling, Blue arrives back to the risk state, which is when the renewal occurs. As soon as it occurs, the risk increases (i.e., Red's effective firing rate increases) and in deterministic time T it eventually reaches the highest risk level. Blue remains in the highest risk level until Blue changes position (i.e., switches to travel state). The decision for Blue is still when to move. We assume the actual time until Blue moves is a random variable (e.g., uniform,

exponential, gamma). The decision variable for Blue is what the parameters of this movement random variable should be. We define

$X_n =$ time between $n-1$ renewal and *n*th renewal
$M_n =$ time until Blue moves for the *n*th time, IID random variables with distribution F
$\delta =$ mean travel time

The travel time distribution does not need to be exponential. All we need for our analysis is the mean: δ. The expected time between renewals is thus

$$E[\mathrm{X}] = \mathrm{E}[\mathrm{M}] + \delta$$

Similar to the CTMC model, Blue may want to spend most of its time in the lower risk portion of the firing time. For concreteness, Blue wants to maximize the time it spends in the α lowest risk portion, that is, between times 0 and αT. In a renewal-reward context, the reward during X_n is the amount of time spent firing between times 0 and αT. Therefore, in this process, the reward is the minimum time of αT and M_n.

$$R_n = \min(\alpha T, M_n) = \begin{cases} \alpha T & if \quad M_n \geq \alpha T \\ M_n & if \quad M_n < \alpha T \end{cases}$$

Then, the expected reward during one renewal period is

$$E(R) = \int_0^{\alpha T} mf(m)dm + (1 - F(\alpha T))\alpha T$$

where *f(m)* is the probability density function (PDF) and *F(M)* is cumulative distribution function (CDF) of distribution of M_n.

Define R_n as the reward accumulated during the *n*th renewal (i.e., time spent in low risk firing during *n*th round of firing). If *R(t)* is the total reward earned up to time *t*, $R(t) = \sum_{n=1}^{N(t)} R_n$, by the renewal-reward limit theorem we can compute the long-run average reward rate as follows.

$$\lim_{t \to \infty} \frac{R(t)}{t} = \frac{E(R)}{E(X)} = \frac{\int_0^{\alpha T} mf(m)dm + (1 - F(\alpha T))\alpha T}{E(M) + \delta}$$

This long-run reward rate is the long-run proportion of time spent firing in the lowest α of the risk spectrum. Therefore, we can obtain the optimal solution that maximizes the long-run average reward rate by determining the best parameter for distribution of M_n.

For example, if $T = 30(\text{min})$, $\delta = 10(\text{min})$, $\alpha = 0.2$ and $M_n \sim \exp(\lambda)$, then

$$\lim_{t\to\infty}\frac{R(t)}{t} = \frac{E(R)}{E(X)} = \frac{\frac{1}{\lambda}(1-(\lambda\alpha T+1)e^{-\lambda\alpha T}) + \alpha T e^{-\lambda\alpha T}}{\frac{1}{\lambda}+\delta}$$

Figure 10 provides the resulting plot. When $\lambda = 0.1545$, it has the highest long-run average reward rate. This is equivalent to moving on average every 6.5 minutes (1/0.1545).

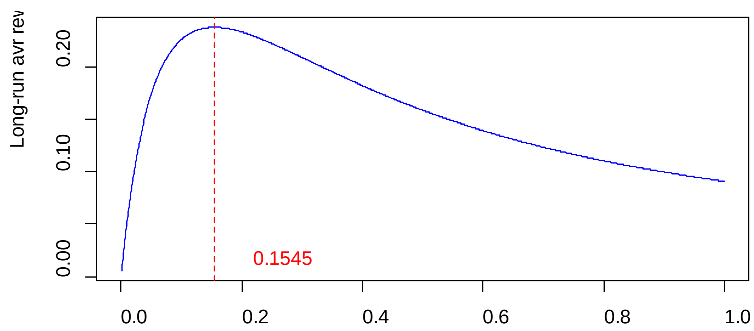


Figure 10. Long-run average reward rate by λ

If we take $\alpha = 1$, which means that Blue accepts all risk levels except for the highest, the optimal λ is 0.058. This is a similar objective to Z_3 in our CTMC model. As shown in Figure 11, this $\alpha = 1$ case for the renewal reward scenario produces the same optimal solution as the CTMC model with the objective function Z_3 for infinite *n*.

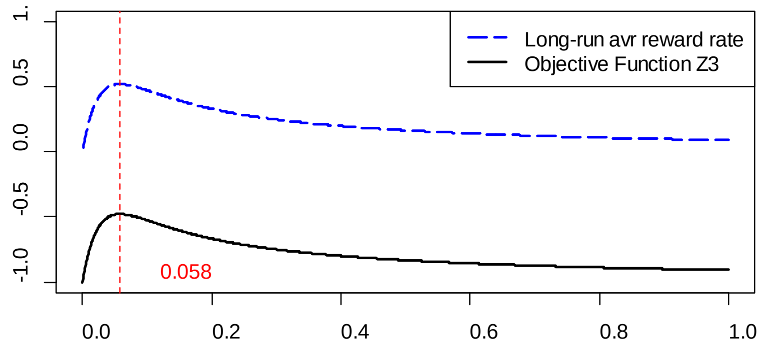


Figure 11. Comparison with the CTMC model

2. Deterministic M

If the μ_i are equal in the original CTMC model, then the time until the system transitions to the highest risk level has an Erlang distribution. In the previous subsection, we assume that the distribution of T is deterministic. Here, we allow any distribution (e.g., uniform, exponential, gamma). The time until Blue moves, M, is deterministic, however, which will be a decision variable in this model. Similarly, with the previous scenario, we define

$X_n =$ time between $n-1$ renewal and nth renewal
$T_n =$ transition time from the lowest risk to the highest risk,
IID random variables with distribution G
$\delta =$ mean travel time

Then, the expected time of X_n should be

$$E[\mathrm{X}] = \mathrm{M} + \delta .$$

Here, we assume the objective is to maximize the time Blue stays at risk levels below the maximum. As discussed at the end of Section B, we assume the commander must choose his move variable M without knowledge of the risk level in real time. Otherwise the commander would move in real time as soon as the risk level hit its maximum. This assumption is valid if the commander has to choose M ahead of time for operational

reasons, or the commander only has an incomplete knowledge of the risk in real time. . The reward is the amount of time in the low risk levels: min(M, T_n).

$$R_n = \min(T_n, M) = \begin{cases} M & if \quad T_n \geq M \\ T_n & if \quad T_n < M \end{cases}$$

Then, the expected reward is

$$E(R) = \int_0^M tg(t)dt + M(1 - G(T))$$

where g*(t)* is the PDF and *G(T)* is the CDF for the distribution of T_n. Then, the long-run average reward rate is

$$\lim_{t\to\infty} \frac{R(t)}{t} = \frac{E(R)}{E(X)} = \frac{\int_0^M tg(t)dt + M(1-G(T))}{M+\delta}$$

Accordingly, we can obtain the optimal solution that maximizes the long-run average reward rate by computing the best deterministic time M. For example, if $\delta = 10(\min)$, $T \sim \exp(\mu = 1/30)$, then

$$\lim_{t\to\infty} \frac{R(t)}{t} = \frac{E(R)}{E(X)} = \frac{\frac{1}{\mu}(1-(\mu M+1)e^{-\mu M}) + Me^{-\mu M}}{M+\delta}$$

As shown in Figure 12, $M = 21.57(\min)$ produces the highest long-run average reward rate.

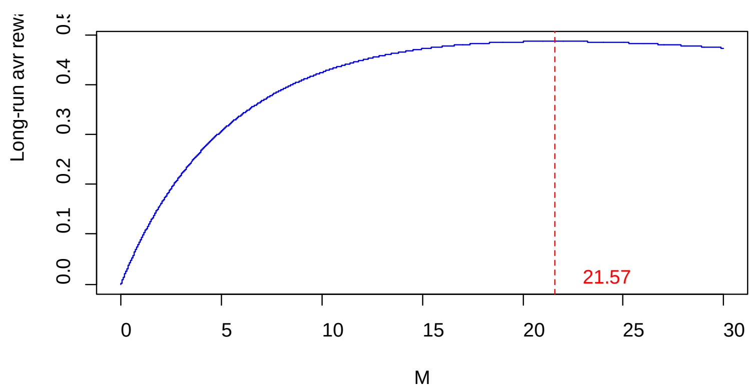


Figure 12. Long-run average reward rate by M ($\delta = 10, \mu = 1/30$)

We can still compute the optimal moving policy even if we relax the assumption that all times have an exponential distribution by using the reward renewal approaches. In addition, the results of this approach are similar to the CTMC model.

H. SUMMARY

We discussed several models and different scenarios in this chapter. In all cases, we use the same travel rate $\delta = 10(\min)$, and the expected time $T = 30(\min)$ to transition from the lowest risk to the highest risk when the expected time until Blue moves $= \infty$. Most of the optimal solutions specify that Blue should move to another position on average every 12 to 24 minutes. Many of the solutions are clustered even more tightly around 15 minutes. There are a few scenarios that produce a much higher rate (see Figure 3), but these result from not adequately penalizing moving and accepting only very low risk states. In conclusion, the results are similar across a variety of modeling assumptions and objectives. If Blue fires at Red with a constant effective firing rate, Blue should spend some amount of time repeatedly firing from the same location, rather than moving to another position immediately after the first fire.

THIS PAGE INTENTIONALLY LEFT BLANK

III. WIN PROBABILITY MODEL WITH TIME LIMIT

In Chapter II, we assume the battle lasts for a long enough time that we can appeal to the limiting distribution of the CTMC and focus on long-run characteristics such as the probability of low risk firing. In reality, the engagement will not last a long time. One side will retreat if it sustains enough damage. Also, the mission may be time critical where it is imperative that Blue forces Red to retreat in a certain time window. If Blue cannot achieve this then Blue effectively “loses” the battle. In this chapter, we carry through the CTMC setup and assumptions from Chapter II. We also incorporate the health of Blue and Red, however, which is directly tied to how many hits each side has received. We also assume Blue has a limited time window to complete its mission. When Blue determines its move policy, it must consider the time and its health. We first define the states of our new CTMC model in Section A. We next describe the model in more detail in Section B. In Section C, we define the probability that Blue wins, which is our primary measure of effectiveness (MOE) when determining a move policy. Subsequently, we propose an optimization algorithm to find the optimal solution to maximize the win probability in Section D. We conclude with numerical examples in Section E.

A. THE STATES

Before developing a model, we need to define the states. Since the model includes three extra components compared to the base model of Chapter II, we represent the state as a vector of four elements:

$$(R, H_B, H_R, T)$$

- R: the risk level
- H_B: the health level of Blue
- H_R: the health level of Red
- T: current time

As in Chapter II, we discretize the state space, so each of the four factors takes on a small number of discrete values. In this chapter, for concreteness we assume each of the four factors takes on four levels, but this is easy to generalize. See Table 2 for a list of the four levels of each factor. “Risk” is treated the same as in Chapter II: there are three risk levels and we include travel as a level in Risk here as travel is effectively the zero-risk level. Each risk level corresponds to a fixed Red effective firing rate, which increases over time when the risk level increases (e.g., because of improved aiming via reaction to Blue fire or intelligence from surveillance). The health status for Red and Blue can either be high, medium, or low. Additional damage after the “Low” level forces the commander to retreat. Moreover, we define some absorbing conditions and divide the absorbing states into two conditions: Win and Lose. If H_B reaches Retreat, then Red “wins” and if H_R reaches Retreat, then Blue “wins.” To model the limited time horizon, we divide that time window into the beginning, middle, and end. After the end state, we assume the battle is over and Blue loses because Blue did not achieve its objective (Red retreat) within the time window. In Section D, we formulate a model to optimize the win probability that the system reaches a Blue win state before a lose state.

Table 2. Status of each component and the number expression

	R	H_B	H_R	T
1	Low	High	High	Begin
2	Mid	Mid	Mid	Mid
3	High	Low	Low	End
4	Travel	**Retreat**	**Retreat**	**Lose**

Bold: absorbing states

For simplicity, we use numbers 1–4 to represent the levels rather than the text in Table 2. For example, the state (1, 2, 3, 1) represents that the risk level is “Low,” the health level of Blue is “Mid,” the health level of Red is “Low” and time is at the beginning of the time horizon. When the system reaches $H_R = 4$ (Red Retreat), Blue wins the battle. When the system reaches $H_B = 4$ (Blue Retreat), Blue loses the battle.

Finally, Blue also loses if the time window closes before Blue forces Red to Retreat; that is, Red wins if $T = 4$ (Lose). Under these assumptions, we define the absorbing states as follows.

- (*, 4, *, *): Blue retreats. Lose states.
- (*, *, 4, *): Red retreats. Win states.
- (*, *, *, 4): Time is over. Lose states.

Mathematically, there are 256 (4×4×4×4) possible states. We can remove several of these states from consideration, however, as multiple absorbing conditions cannot occur simultaneously. For example, Blue and Red cannot retreat at the same time or Blue cannot retreat after the time window has closed. Moreover, since we assume that Blue and Red cannot be hit by the enemy while they move, it is not possible that a moving state (i.e., (4, *, *, *)) has any absorbing condition. As a result, the number of possible states decreases to 198 after we remove the impossible 58 states. The final valid state space consists of 108 (4×3×3×3) transient states and 90 (3×1×3×3 + 3×3×1×3 + 4×3×3×1) absorbing states.

B. MODEL DESCRIPTION

The system starts at a state (1, 1, 1, 1) that represents low risk and high health of Blue and Red in the beginning of the battle. After some amount of firing time, the system transitions to one of the following five states, depending on what happens first. We provide more detail about each of these five state changes as follows.

i.	Risk level increases:	(2, 1, 1, 1)
ii.	Blue’s health decreases:	(1, 2, 1, 1)
iii.	Red’s health decreases:	(1, 1, 2, 1)
iv.	Time horizon changes:	(1, 1, 1, 2)
v.	Blue moves:	(4, 1, 1, 1)

i) The risk level gradually increases over time as Red obtains better information about Blue's location. This information may come from radar signals that track Blue's fire or surveillance assets, such as UAVs. The time until the risk increases by one level is exponentially distributed with μ_R. As discussed in Chapter II, this time includes, for example, the time required to process intelligence to determine an updated aimpoint and the time to recalibrate and aim the artillery for the new aimpoint. For simplicity, we assume the rate μ_R does not depend on the current risk level, although this is easy to generalize. Therefore, the state (1, 1, 1, 1) transitions to (2, 1, 1, 1) with rate μ_R.

ii) & iii) The health status for Blue and Red decrease over time since they fire at each other continuously until one of them retreats or the time window closes. The time until the health level of Blue (Red) changes has an exponential distribution with rate $\mu_{HB}(j)$ ($\mu_{HR}(j)$), where the parameter *j* dictates the current risk level. The health rate of Blue (Red) corresponds to the effective firing rate of Red (Blue). That is Red's effective fire in risk level *j* is a Poisson process with rate $\mu_{HB}(j)$. Recall Red's effective firing rate corresponds to the overall rate Red fires rounds multiplied by the probability a round hits Blue. Rather than defining one parameter for overall firing rate and one parameter for hit probability, for simplicity we just define Red's effective firing rate, which corresponds to Blue's health rate $\mu_{HB}(j)$. These assumptions imply that one hit from Red fire decreases the health of Blue by one level. If multiple hits are required to decrease the health by one level, then we can, for example, define $\mu_{HB}(j)$ as the Red effective firing rate divided by the number of hits per health level. A higher effective firing rate results in faster reduction of the other's health. In higher risk states, the rate $\mu_{HB}(j)$ will be higher than for lower risk levels. As a result, the state (1, 1, 1, 1) can transit to (1, 2, 1, 1) with rate $\mu_{HB}(1)$ or (1, 1, 2, 1) with rate $\mu_{HR}(1)$. We treat health transitions as independent of risk transitions. A health transition corresponds to a direct hit by either side, which translates into potentially useful intelligence, which could result in an increased effective firing rate (and hence an increased risk level). Therefore, one could model a hit as changing both the

health level and risk level simultaneously. We do not model the system evolution in this fashion, but leave it as a suggestion for future work.

iv) Eventually, the time window will close. We divide the time window into three levels; the time until the time component changes has an exponential distribution with mean $1/\mu_T$. That is, the time component level increases from level *j* to *j+l* according to an exponential distribution with rate μ_T. The rate parameter μ_T does not depend upon the risk level or health status of Red or Blue. For example, if the desired time window is 60 minutes, then $\mu_T = 1/20(\text{min})$. With this assumption, the time until the window closes has a Gamma distribution with shape parameter 3 and rate parameter μ_T. As discussed in Chapter II, the finer we divide the time window into levels, the more deterministic it becomes. The computational complexity also grows significantly, however. The state (1, 1, 1, 1) transitions to (1, 1, 1, 2) with rate μ_T.

v) Blue moves to avoid high risk levels. The time until Blue moves has an exponential distribution with rate λ. The rate λ is a decision variable for Blue. In Chapter II, λ was a scalar. In this chapter, we still assume that Blue does not know the risk level and we also assume Blue does not know Red's health status. Blue knows its own health status and Blue knows the current time, however. Thus, λ is a function of Blue health status and time. For example, if Blue's health is low and time is in the beginning, Blue may want to move frequently to avoid the enemy's fire. On the other hand, if Blue's health is high and time is at the end, Blue may want keep firing without moving to increase the chances of forcing Red to retreat. We will discuss this in more detail in Section D. As an example, the state (1, 1, 1, 1) transitions to (4, 1, 1, 1) with rate λ.

Figure 13 illustrates the five possible transitions. It is only possible to have all five transition types when the system is in a "Low" or "Mid" risk transient state.

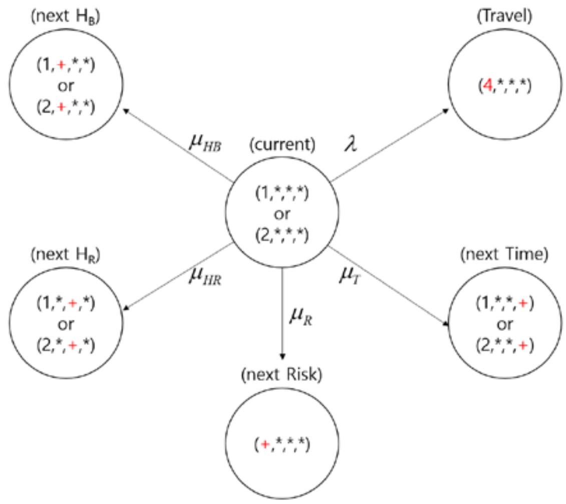


Figure 13. Transition diagram when the system is in "Low" or "Mid" risk transient states

When the system is in the "High" risk level (i.e., (3, *, *, *)), there is no possible next risk level since we only have three levels of risk. Figure 14 shows the transition diagram reflecting these states.

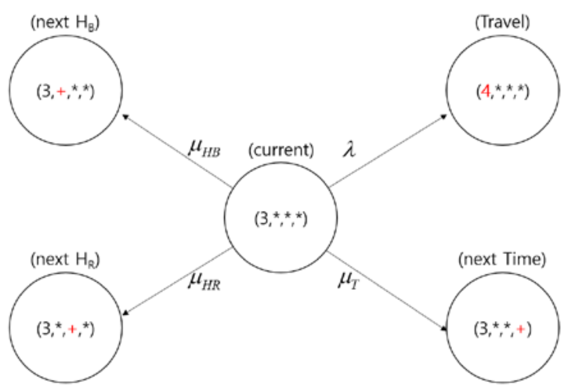


Figure 14. Transition diagram when the system is in "High" risk transient states

In addition, when the system is in "Travel" states, it can only go to the lowest risk level state or the next time horizon state. Figure 15 shows the transition diagram for these states.

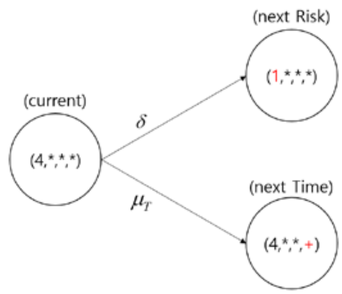


Figure 15. Transition diagram when the system is in "Travel" transient states

The system only evolves in one way for health and time. Eventually, the system reaches one of the absorbing states, Win or Lose states, and the transitions stop.

C. WINNING PROBABILITY

To compute the win probability, we need to compute the probability the system next transitions to each state. Suppose that T_1 and T_2 are independent exponential random variables with rate μ_1 and μ_2 respectively. Then the probability that T_1 is less than T_2 is

$$P(T_1 < T_2) = \frac{\mu_1}{\mu_1 + \mu_2}$$

Likewise, if $T_1, \ldots, T_n$ are independent exponential random variables with rates $\mu_1, \ldots, \mu_n$ respectively, then the probability that T_1 is smaller than the others (i.e., T_1 is the minimum) is

$$P(T_1 = \min(T_1, \ldots, T_n)) = \frac{\mu_1}{\sum_{i=1}^{n} \mu_i}$$

See chapter 5 of Ross (2014) for details.

For example, in our model, the probability that the system transitions from state (1, 1, 1, 1) to state (1, 1, 1, 2) is

$$P\{(1,1,1,1) \to (1,1,1,2)\} = \frac{\mu_T}{\mu_R + \mu_{HB}(1) + \mu_{HR}(1) + \mu_T + \lambda_{(1,1,1,1)}}$$

and the probability that the system transitions to state (4, 1, 1, 1) is

$$P\{(1,1,1,1) \to (4,1,1,1)\} = \frac{\lambda_{(1,1,1,1)}}{\mu_R + \mu_{HB}(1) + \mu_{HR}(1) + \mu_T + \lambda_{(1,1,1,1)}}$$

where $\lambda_{(1,1,1,1)}$ is the moving rate in state (1, 1, 1, 1). We define the probability that the system transitions from state *i* to state *j* as P_{ij}, which is the standard Markov transition probability.

Blue chooses its move strategy (i.e., its λ vector: $(\lambda_{(1,1,1,1)}, \lambda_{(1,1,1,2)}, \cdots, \lambda_{(3,3,3,3)})$) to maximize the probability that Blue wins. We compute this win probability for a fixed λ vector in this section. To compute the probability, we set up a system of equations. In this system, we assign each state to one of three categories: Blue Win, Blue Lose, or "Neutral." The Neutral category denotes that the battle is still ongoing. $P(s)$ is the probability Blue eventually wins, given the system is currently in state *s*. Thus, $P(s)=1$ if *s* is a win state and $P(s)=0$ if *s* is a lose state. If *s* is a Neutral, however, we need to consider the next state and condition on it. Using the Law of Total Probability, we have

$$P(s) = \sum_{i \in states} P_{si} \times P[win \mid s \to i] \quad \forall s \in Neutral$$

where P_{si} is the transition probability from state *s* to state *i*. By the Markov property, we can say that $P[win \mid s \to i] = P(i)$. Thus, for all neutral states *s* we have

$$\begin{aligned} P(s) &= \sum_{i \in states} P_{si} \times P(i) \\ &= \sum_{i \in Win} P_{si} + \sum_{i \in Neutral} P_{si} \times P(i) \quad \forall s \in Neutral \end{aligned}$$

We can write out a matrix equation for this system of equations.

$$P_{win} = P_{n \to w} + P_D \times P_{win}$$

where P_{win} is a vector containing $P(s)$ for all neutral states s and P_D is the part of the transition matrix that just tracks transitions among neutral states. Also, $P_{n\to w}$ is a vector representing the probability of transitioning from a given neutral state to a win state in one transition. We can solve for the desired win probabilities using the following formula:

$$P_{win} = (I - P_D)^{-1} P_{n\to w}$$

The vector P_{win} is a function of the state s, but it is also a function of all the other parameters, such as the rates of the exponential distributions. Most importantly for our purposes, P_{win} depends upon the λ vector. Our goal is to determine the optimal move policy via the λ vector to maximize the win probability.

D. OPTIMIZATION

Blue may want to consider an objective that accounts for $P(s)$ for multiple states s. We focus on the first state, (1, 1, 1, 1), however, because we assume that the engagement starts with 100% health and enough time to conduct a mission at time 0. Denote *P(1, 1, 1, 1)* as the probability of Blue winning starting in state (1, 1, 1, 1), that is, at the beginning of the battle. In order to maximize the probability *P(1, 1, 1, 1)*, as mentioned previously, we must solve for the optimal λ vector. It is realistic that Blue knows its health level and the time spent in battle. In other words, we assume Blue can vary the moving rate λ depending on its current health level and the current time window. If some states have the same Blue health level and time window, the optimal moving rate λ should be the same in those states even if the risk level and Red health level vary. For instance, Blue may want to move infrequently and spend more in firing states (i.e., risk levels 1, 2, 3) when Blue has a high health level and limited time because these are the only states that can decrease Red's health. On the other hand, Blue may want to move quickly when Blue's health level is low even though the risk to Blue may be low because this limits the possibility that Blue's health will decrease further.

Because Blue only accounts for its health and the time when choosing the move strategy, there are 9 possible moving rates. We group these rates in a vector

$\lambda^T = (\lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{13}, \lambda_{22}, \lambda_{31}, \lambda_{23}, \lambda_{32}, \lambda_{33})$ where λ_{ij} is the moving rate in states (*, *i*, *, *j*) for all $i = 1,2,3$ and $j = 1,2,3$. We use an iterative backward method to find the optimal vector of rate λ. We start with $i = 3$, $j = 3$ and compute λ_{33}. If we knew which particular (*, 3, *, 3) state we first transition into (e.g., (1, 3, 2, 3)), then we could compute the probability of Blue winning starting from that (*, 3, *, 3) state using the approach from Section C, and optimize with respect to λ_{33}. Unfortunately, there are 12 possible (*, 3, *, 3) states: (1, 3, 1, 3), (1, 3, 2, 3), (1, 3, 3, 3), …, (4, 3, 1, 3), (4, 3, 2, 3) and (4, 3, 3, 3). We do not know which of these 12 (*, 3, *, 3) states the system will first transition to starting from (1, 1, 1, 1). Using a similar approach to Section C, however, we can compute the probability that starting from (1, 1, 1, 1) the system will first transition to, for example, (4, 3, 1, 3) out of all the (*, 3, *, 3) states. Using these first-passage probabilities, we assign a weight to each (*, 3, *, 3) state. For example, let *u* be a particular (*, 3, *, 3) state. Then we set the weight in state *u* as follows.

$$weight(u) = \frac{P\left[state(1,1,1,1) \rightarrow state(u)\right]}{\sum_{v \in (*,3,*,3)} P\left[state(1,1,1,1) \rightarrow state(v)\right]}, \quad \forall u \in (*,3,*3)$$

Consequently, the optimization problem becomes

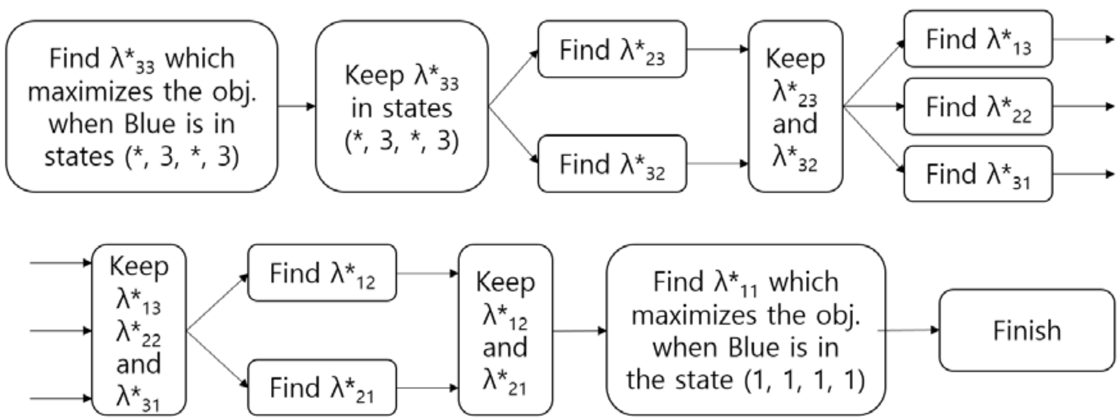
$$\max_{\lambda_{33}} \quad z = \sum_{u \in (*,3,*,3)} weight(u) \times P(u)$$

where *P(u)* is the probability to win given the system is in state *u*. To compute *weight(u)* requires knowledge of all λ_{ij}, which is what we are trying to compute. This is where the iterative aspect of the algorithm comes into play. In the first round, we initialize all $\lambda_{ij} = 0$. This allows us to compute *weight(u)* for all (*, 3, *, 3) states, and hence optimize λ_{33}. Once we have λ^*_{33}, we determine the optimal rate λ^*_{23} in states (*, 2, *, 3) by computing *P(u)* and *weight(u)* for all (*, 2, *, 3) states. We compute *P(u)* by using the λ^*_{33} computed earlier, and we compute *weight(u)* by assuming all other $\lambda_{ij} = 0$. We

continue working our way backwards in this fashion to determine λ^*_{ij} for all $i = 1, 2, 3$ and $j = 1, 2, 3$:

$$\max_{\lambda_{ij}} \quad z = \sum_{u \in (*, i, *, j)} weight(u) \times P(u)$$

where λ_{ij} is the moving rate in states (*, *i*, *, *j*) and *P(u)* is the probability to win when the system is in state *u*. We illustrate the algorithm in Figure 16.



λ^*_{ij} : optimal moving rate in states (*, *i*, *, *j*)

Figure 16. The optimization algorithm

To compute the optimal rate λ^*_{11}, we do not weight all (*, 1, *, 1) states; we know the system starts in state (1, 1, 1, 1). Thus, we optimize λ^*_{11} with respect to only state (1, 1, 1, 1) by computing *P(1, 1, 1, 1)* directly.

After one round, we have estimates λ^*_{ij}. We compute these rates by using *weight(u)* derived from assuming $\lambda_{ij} = 0$, however. In round 2 we compute the *weight(u)* by using the λ^*_{ij} calculated in round 1, which generates new estimates of λ^*_{ij} in round 2. We continue this iterative approach until the optimal vector λ^* converges.

E. NUMERICAL DEMONSTRATION

To implement this model, we use the R: A language and environment for statistical computing (R core team, 2016). We use the `optimize()` function to find the optimal solutions in R. At any time, we are only ever computing one optimal parameter so the optimization is straightforward. First, we discuss the values of the parameters we use in this section and then demonstrate the algorithm.

1. Parameters

For the purpose of comparison with the numerical demonstration result of the long-run risk model in Chapter II, we use the same travel rate δ and expected time T to transition from low risk to high risk, when Blue does not move (i.e., $\lambda = 0$) as follows:

$$\delta = \frac{1}{10}, \qquad T = 30$$

Consequently, we use the risk transition rate $\mu_R = \frac{1}{15}$ to keep the time $T = 30$. We add a constraint to make our problem more realistic. Blue has to fire at least one shot after arriving at a new position since Blue's primary purpose is firing at Red. We enforce this by placing a maximum value on λ_{ij}. In the following experiments, we assume

$$\lambda_{\max} = \frac{1}{5}$$

which implies it requires on average at least 5 minutes to fire one shot (mission) and be ready to move. For the health transition rates $\mu_{HB}(j)$ and $\mu_{HR}(j)$, we assume the expected time to transition from the high health status to retreat, when there is no Blue movement ($\lambda = 0$), is the same for both Blue and Red. We assume the specific μ values differ for each side, however. Blue has a lower health transition rate than Red in the low risk state but a higher health transition rate in the high risk state. This is reasonable because we assume that Red does not know Blue's location before Blue fires in a new position; hence, in the low risk state, Red has little chance to decrease Blue's health. As

Blue fires, however, Red's threats increase as Red pinpoints Blue's location. In this example, Blue's accuracy does not improve when it shoots successively from the same location. With this assumption, we use the following parameters.

$$\mu_{HB}(1) = \frac{1}{30}, \qquad \mu_{HB}(2) = \frac{1}{20}, \qquad \mu_{HB}(3) = \frac{1}{10}$$

$$\mu_{HR}(j) = \frac{1}{20}, \quad \forall j = 1, 2, 3$$

Lastly, we assume the limited time is 2 hours (120 minutes), which makes the rate μ_T as follows.

$$\mu_T = \frac{1}{time\,limit/3} = \frac{1}{40}$$

2. Algorithms

To compute the optimal solutions, we initialize our vector λ to the 0 vector. This represents the situation when Blue does not move. We next compute the optimal rate λ^*_{33}. Thus, the objective function for the rate λ_{33} is

$$\max_{\lambda_{33}} \quad z = \sum_{u \in (*,3,*,3)} weight(u) \times P(u)$$

We use the approach described in Section D to solve for λ^*. In the first round of the algorithm we initialize $\lambda = 0$. The *weight(u)* and *P(u)* in (*, 3, *, 3) states at the beginning of the algorithm when all $\lambda = 0$ appear in Table 3. Figure 17 displays the updated λ_{33} value, which we compute in the first part of round 1.

Table 3. The weights and win probabilities in (*, 3, *, 3) states when $\lambda = 0$

State: *u*	(1,3,1,3)	(1,3,2,3)	(1,3,3,3)	(2,3,1,3)	(2,3,2,3)	(2,3,3,3)
weight(u)	0.0294	0.0336	0.0240	0.0607	0.0827	0.0677
P(u)	0.0631	0.1675	0.4230	0.0400	0.1224	0.3602
State: *u*	(3,3,1,3)	(3,3,2,3)	(3,3,3,3)	(4,3,1,3)	(4,3,2,3)	(4,3,3,3)
weight(u)	0.1634	0.2728	0.2659	0	0	0
P(u)	0.0233	0.0816	0.2857	0.0505	0.1340	0.3384

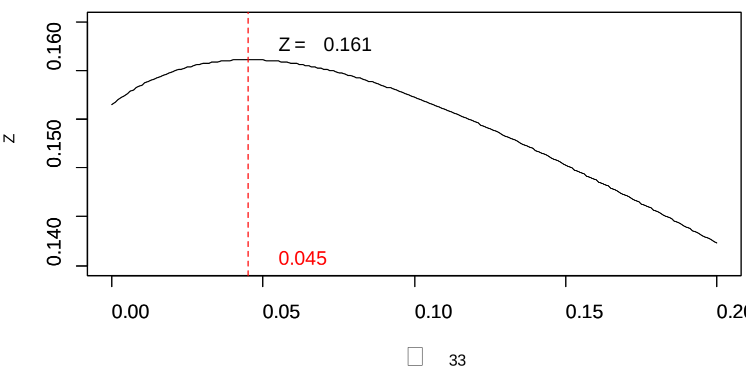


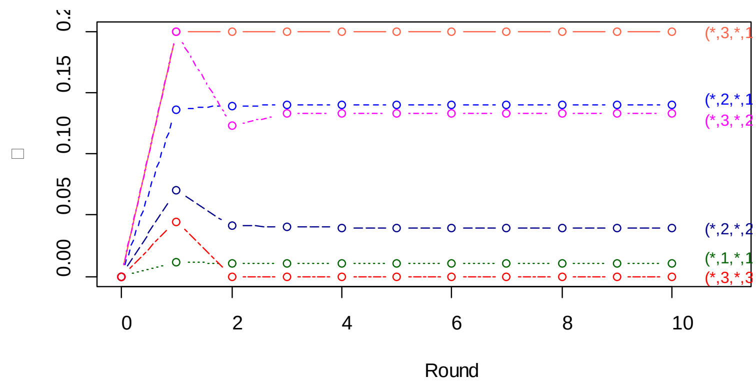
Figure 17. The solution for λ_{33} in round 1

We replace the rate $\lambda_{33} = 0$ in (*, 3, *, 3) states with the updated rate $\lambda^*_{33} = 0.045$ from the first round. With the same repetitive method, following the algorithm described earlier provides us with the values of the updated vector λ^* in round "1" (Table 4).

Table 4. The updated vector λ^* after round "1"

λ^*_{ij}	λ^*_{11}	λ^*_{12}	λ^*_{21}	λ^*_{13}	λ^*_{22}	λ^*_{31}	λ^*_{23}	λ^*_{32}	λ^*_{33}
value	0.0118	0	0.1367	0	0.0703	0.2	0.2	0	0.0449

We iterate this algorithm updating the vector λ^* until the maximum absolute difference between the λ vector in one round and the previous round is less than 10^{-6}. The results are shown in Figure 18 and Figure 19.



The 6 sets of family states have positive rate λ_{ij} but others have zero in all iterations. The rates have no big variation after round 3. It converges at round 7.

Figure 18. The changes of the rate λ_{ij} during 10 rounds

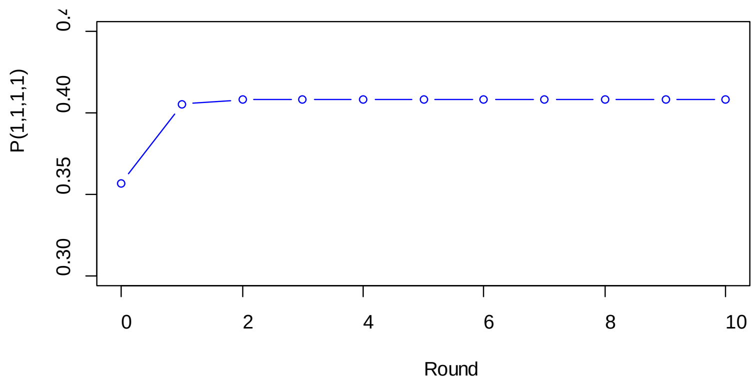


Figure 19. The changes of the objective value: *P(1,1,1,1)* during 10 rounds

After two rounds, we have a near optimal solution and it converges at round 7. The probability Blue wins increases from 0.357 with no moving to 0.408 by taking the optimal vector λ^*. Utilizing this model provides Blue with a 5% greater chance to win than if Blue uses a stationary artillery approach without moving. The probability Blue wins is less than 0.5 because if the time window closes, Blue loses. The final converged values of the optimal vector λ^* appear in Table 5.

Table 5. The converged optimal vector λ^*

λ^*_{ij}	λ^*_{11}	λ^*_{12}	λ^*_{21}	λ^*_{13}	λ^*_{22}	λ^*_{31}	λ^*_{23}	λ^*_{32}	λ^*_{33}
value	0.0108	0	0.1406	0	0.04	0.2	0	0.1335	0

When the time window closes soon (i.e., in state (*, *, *, 3)), Blue should not move ($\lambda = 0$). By staying in the same position Blue can achieve a higher firing rate, which increases Blue's win probability when time is running out. On the other hand, when Blue has a lower health but (probabilistically) plenty of time (i.e., in states (*, 2, *, 1), (*, 3, *, 1) and (*, 3, *, 2)), Blue should move frequently. The optimal rates λ_{ij} for these states lie in the interval (0.1335, 0.2), which represents that Blue moves on average every (5, 7.5) minutes. By moving frequently Blue can avoid Red's shells, which decreases Blue's firing rate. Survivability is more important for Blue in this situation, however. Especially, in state (*, 3, *, 1), where Blue has the maximum moving rate (i.e., ${\lambda^*}_{31} = \lambda_{\max} = 0.2$), Blue should move very frequently as there is little to gain for Blue by exposing itself to more risk early in the battle. In addition, when Blue is at its maximum health (i.e., in state (*, 1, *, *)), Blue should move very infrequently. The time window component has a negligible impact on this result. The optimal rates λ_{ij} are distributed in (0, 0.0108), which implies that Blue moves on average every (92.6, ∞) minutes. Blue has a higher health and thus can endure some risk for the benefit of a higher firing rate in the same position.

We conclude this chapter by examining the situation where Blue's accuracy increases at higher risk levels. Instead of the constant health transition rate for Red, $\mu_{HR}(j) = 1/20$, Red transitions quickly to "Retreat" in the higher risk level. With this additional assumption, the health transition rates are

$$\mu_{HB}(1) = \frac{1}{30}, \qquad \mu_{HB}(2) = \frac{1}{20}, \qquad \mu_{HB}(3) = \frac{1}{10}$$

$$\mu_{HR}(1) = \frac{1}{25}, \qquad \mu_{HR}(2) = \frac{1}{20}, \qquad \mu_{HR}(3) = \frac{1}{15}$$

We still hold other assumptions, however, that i) Blue has a lower health transition rate than Red in the low risk state and a higher health transition rate in the high risk state, and ii) the expected time from the high health state to the retreat state is the same. Table 6

shows the results. Blue wins with probability 0.399 in this scenario, which is very similar to the original example.

Table 6. A result with increasing $\mu_{HR}(j)$

λ^*_{ij}	λ^*_{11}	λ^*_{12}	λ^*_{21}	λ^*_{13}	λ^*_{22}	λ^*_{31}	λ^*_{23}	λ^*_{32}	λ^*_{33}
value	0	0	0.0734	0	0	0.1999	0	0.0468	0

Blue does not move in 6 of the 9 categories. Blue moves frequently only in (*, 3, *, 1) states. Even if Blue's health is low (i.e., in states (*, 3, *, 2) and (*, 3, *, 3)), Blue stays in the same position and continues firing at Red to achieve a higher accuracy. If Blue has an ability to increase its accuracy during firing, for example adjusting aims with some feedback, Blue should spend more time firing before moving to another position.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. CONCLUSION

A. SUMMARY

In this thesis, we describe two models to analyze shoot-and-scoot policies for artillery forces. The shoot-and-scoot tactic is important, but there appears to be limited quantitative analysis on the move decision. Currently, commanders use their experience and intuition to determine when the artillery should change locations. Most commanders are risk averse, so they tend to move frequently to avoid the enemy's counter-fire. Frequently moving limits the potential benefits of a higher firing rate and improved accuracy.

A primary component of our models is "risk," which increases over time when Blue stays in the same position. In Chapter II, we develop a long-run risk model, which only considers risk and assumes the battle goes on for a long period of time. We examine several different objective functions that consider both risk and firing rate. The main objective of this model is to limit Blue's exposure to higher risk. In Chapter III, we construct the win-probability model in a limited time window scenario. In this model, we incorporate other factors such as "Health" and "Time in battle." The battle does not go on for an arbitrarily long time, but instead Blue must win within a finite time window. The objective of this model is maximizing the probability Blue wins. The decision variables in both models are the rates at which Blue moves.

Although we examine only one representative scenario in each model, the parameters are reasonable according to the author's experience. The general result is that in most situations Blue should spend a reasonable amount of time engaging with Red in artillery fire from the same location. When we account for time and health (Chapter III), this result becomes even more pronounced. Blue should never move in certain states (e.g., high Blue health, later in the battle). Moving frequently reduces risk to Blue, but limits Blue's ability to inflict damage on Red. This result may run counter to the approach of some commanders, who believe they should move frequently to survive and

win the battle. Our result should provide the commanders with some insight about shoot-and-scoot tactics.

Another contribution of this thesis is that the models can offer a method to evaluate the best strategy for the artillery forces. The parameters we use in this thesis may not be realistic in all scenarios. Our models are transparent and straightforward, however, so users can input their own parameters based on, for example, estimates using real battle data.

B. FUTURE WORK

In our model, time is the metric for risk. Red's firing rate increases with time as Red obtains more intelligence about Blue's location. Implicitly we assume this occurs primarily as Red reacts to Blue’s artillery fire. However, it may also increase in time for other reasons such as surveillance reports from UAVs. Future work could model risk as being explicitly connected to the number of Blue rounds fired rather than just time. Currently we have one measure of risk, which increases in time. We could model risk to Blue ("Blue risk") and risk to Red ("Red risk") separately. Blue would want to be in low Blue risk states and high Red risk states, which correspond to a high relative effective firing rate. Finally, in the win-probability model, the risk and health levels evolve independently. In reality both are tied to accurate fire, so future work could model the interaction between health and risk.

One of the limitations of this thesis is that we consider mainly exponential distributions in order to leverage Markov models. In many real situations, the exponential may not be realistic. Although we use reward renewal process approaches to use other distributions, more general methods could be used to look at other distributions. We suggest a simulation model be developed, which would allow great flexibility for probability distributions and finer resolution of modeling detail. It would be interesting to compare the results from simulation analysis to our model.

Another limitation is that we explore just one scenario in each model. This limits our ability to generalize our insights. Future work could perform more rigorous sensitivity analysis, perhaps taking a design of experiments approach. This would

generate more general insights about shoot-and-scoot tactics. For example, a study that varies the expected time it takes to move positions would provide insight into how much training should be done to potentially reduce the time required to move. There are numerous possible scenarios to analyze and the results would offer the effective strategy recommendations for artillery forces.

Future work could incorporate more complicated, but realistic, aspects. Examples include feedback or reinforcements. For example, Blue may receive better feedback about its aimpoint accuracy when the assets that provide information about the target and impact points (e.g., surveillance UAVs) can operate effectively. If these support assets can operate freely close to Red, then Blue’s accuracy can increase quickly. If Red takes measures to eliminate those assets, however, then Blue’s accuracy may not improve much by staying at the same location. If Blue can receive reinforcements, then it is possible that Blue’s health could increase during the course of the battle. Currently in Chapter III, Blue’s health only decreases.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Calvin, W. S. (1971). A single-shot noisy duel with detection uncertainty. *Operations Research, 19*(1):170–181. Retrieved from http://dx.doi.org/10.1287/opre.19.1.170

Christy, D. E. (1969). *A Lanchester based model for analyzing infantry fire and maneuver tactics* (Master's thesis). Retrieved from Calhoun http://hdl.handle.net/10945/12536

Duke, J. R. (1996). *A Markov model for parametric sensitivity analysis of Crusader effectiveness* (Master's thesis). Retrieved from Calhoun http://hdl.handle.net/10945/32072

Gautam, K. P. (2010). *Artillery lessons from the early 21st century*. Retrieved from http://www.claws.in/images/publication_pdf/1272003620IB-16.__22.04.10.pdf

Harari, L. (2008). *Defense against rocket attacks in the presence of false cues* (Master's thesis). Retrieved from Calhoun http://hdl.handle.net/10945/3870

Holmes, R., Strachan, H., Bellamy, C., & Bicheno, H. (2001). *The Oxford companion to military history*. New York: Oxford University Press.

K9 Thunder self-propelled howitzer. (2014, July 1). Retrieved from http://www.militaryfactory.com/armor/detail.asp?armor_id=484

Koba, G. M. (1996). *Artillery strike force* [Monograph]. Retrieved from Defense Technical Information Center http://www.dtic.mil/docs/citations/ADA324363

Kress, M. (1991). A two-on-one stochastic duel with maneuvering and fire allocation tactics. *Naval Research Logistics (NRL)*, *38*(3), 303–313. Retrieved from http://dx.doi.org/10.1002/1520-6750(199106)38:3<303::AID-NAV3220380303>3.0.CO;2-F

Park, S. (2015). *An analysis on the artillery tactics with absorbing Markov chain* (Master's thesis). Korea National Defense University, Republic of Korea.

R Core Team (2016). *R: A language and environment for statistical computing* [Computer software]. R foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Ravid, I. (1989). Defense before or after Bomb-Release-Line. *Operations Research 37*(5): 700–715. Retrieved from http://dx.doi.org/10.1287/opre.37.5.700

Ross, S. M. (2014). *Introduction to probability models* (11th ed.). Amsterdam; Boston: Academic Press.

Sweat, C. W. (1971). A single-shot noisy duel with detection uncertainty. *Operations Research 19*(1):170-181. Retrieved from http://dx.doi.org/10.1287/opre.19.1.170

U.S. Marine Corps. (2002). *Tactics, techniques, and procedures for fire support for the combined arms commander* (FM 3–09.31 MCRP 3–16C). Washington, DC: Author.

Washburn, A. R. (2002). *Notes on firing theory* [Lecture notes]. Retrieved from Naval Postgraduate School website http://faculty.nps.edu/awashburn/Files/Notes/FiringTheory.pdf

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California